

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
17 February 2005 (17.02.2005)

PCT

(10) International Publication Number  
**WO 2005/014796 A2**

Document AM24  
Appl. No. 09/517,466

- (51) International Patent Classification<sup>7</sup>: C12N
- (21) International Application Number: PCT/US2004/025831
- (22) International Filing Date: 9 August 2004 (09.08.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/493,322 8 August 2003 (08.08.2003) US
- (71) Applicant (for all designated States except US): INVIT-ROGEN CORPORATION [US/US]; 1600 Faraday Avenue, Carlsbad, California 92008 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): CHESNUT, Jonathan D. [US/US]; 1839 Valencia Avenue, Carlsbad, California 92008 (US). MADDEN, Knut R. [US/US]; 2705 Glasgow Drive, Carlsbad, California 92008 (US). DUDAS, Miroslav [US/US]; 276 Trilogy Street, San Marcos, California 92078 (US). LEONG, Louis [SG/US]; 24778 Butler Road, Junction City, OR 97448 (US). HARRIS, Adam N. [US/US]; 5040 Codorniz Way, Oceanside, California 92057 (US).
- (74) Agents: ESMOND, ROBERT W.? et al.; STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C., 1100 New York Avenue, N.W., Washington, District of Columbia 20005 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS AND COMPOSITIONS FOR SEAMLESS CLONING OF NUCLEIC ACID MOLECULES

(57) Abstract: The present invention is in the fields of biotechnology and molecular biology. More particularly, the present invention relates to cloning or subcloning one or more nucleic acid molecules comprising one or more type II restriction enzyme recognition sites. The present invention also embodies cloning such nucleic acid molecules using recombinational cloning methods such as those employing recombination sites and recombination proteins. The present invention also relates to nucleic acid molecules (including RNA and iRNA), as well as proteins, expressed from host cells produced using the methods of the present invention.

WO 2005/014796 A2

# METHODS AND COMPOSITIONS FOR SEAMLESS CLONING OF NUCLEIC ACID MOLECULES

## Background of the Invention

### Field of the Invention

[0001] The present invention is in the fields of biotechnology and molecular biology. More particularly, the present invention relates to seamlessly cloning or subcloning one or more nucleic acid molecules. The present invention also relates to seamless cloning of nucleic acid molecules comprising one or more type II restriction enzyme recognition sites. The present invention also embodies cloning such nucleic acid molecules using recombinational cloning methods such as those employing recombination sites and recombination proteins. The present invention also relates to nucleic acid molecules (including RNA and iRNA), as well as proteins, expressed from host cells produced using the methods of the present invention.

### Related Art

[0002] A significant problem with many of the currently available molecular cloning techniques results from the reliance upon restriction sites. These techniques result in the presence of extraneous polynucleotides in the amplification products even after restriction digestions. Such extraneous polynucleotides can introduce design limitations on the cloned product which often interfere with the structure and function of the desired gene products, be they RNA, DNA or protein.

[0003] One method of joining nucleic acids without introducing extraneous bases or relying on the presence of restriction sites is splice overlap extension (SOE) (Yon *et al.*, *Nucl. Acids Res.* 17:4895 (1989) and Horton *et al.*, *Gene* 77:61-68 (1989)). This method is based on the hybridization of homologous 3' single-stranded overhangs to prime synthesis of DNA using each complementary strand as a template. Although this technique can join fragments without introducing extraneous nucleotides (in other words,

seamlessly), it does not permit the easy insertion of a DNA segment into a specific location when seamless junctions at both ends of the segment are required. Nor does this technique allow for joining fragments with a vector. Ligation with a vector must be subsequently performed by incorporating restriction sites onto the termini of the final SOE fragment. Finally, this technique is particularly awkward when trying to exchange polynucleotides encoding various domains or mutation sites between genetic constructs encoding related proteins.

[0004]       Sorge *et al.*, U.S. Patent No. 6,261,797 describe a method by which polynucleotide sequences of interest are synthesized using one or more synthesis primers, wherein at least one of the primers is a releasable primer. After synthesis, the synthesis product is cleaved by a releasing enzyme. The releasable primers of Sorge *et al.* comprise a recognition site for a type IIs restriction endonuclease, principally *Eam1105I*. This then allows for "seamless domain replacement" where synthesis reactions allow the production of a polynucleotide of interest by synthesizing two different polynucleotide sequences using separate sets of primers, cleaving the synthesis products with a releasing enzyme, and ligating together the two sets of release synthesis products.

#### Type IIs Restriction Enzymes

[0005]       Restriction enzymes can be grouped based on similar characteristics. In general there are three major types or classes: I, II (including IIs) and III. Class I enzymes cut at a somewhat random site from the enzyme recognition sites (see Old and Primrose, Principles of Gene Manipulation, Blackwell Sciences, Inc., Cambridge, Mass., (1994)). Most enzymes used in molecular biology are type II enzymes. These enzymes recognize a particular target sequence (*i.e.*, restriction endonuclease recognition site) and break the polynucleotide chains within or near to the recognition site. The type II recognition sequences are continuous or interrupted. Class IIs enzymes (*i.e.*, type IIs enzymes) have asymmetric recognition sequences. Cleavage occurs at

a distance from the recognition site. These enzymes have been reviewed by Szybalski *et al.* *Gene* 100:13-26 (1991). Class III restriction enzymes are rare and are not commonly used in molecular biology.

[0006] Type-IIs endonucleases generally recognize non-palindromic sequences and cleave outside of their recognition site, thus producing overhangs of ambiguous base pairs. (Szybalski, *Gene* 40:169-173 (1985).) Additionally, as a result of their non-palindromic recognition sequences, the use of type-IIs endonucleases will generate more markers per kB than a similar type-II endonuclease, e.g., approximately twice as often. U.S. Patent No. 4,293,652 discloses a linker with a type-IIs enzyme recognition sequence to permit synthesized DNA to be inserted into a vector without disturbing a recognition sequence. Brousseau *et al.* (*Gene* 17:279-289 (1982)) and Urdea *et al.* (*Proc. Natl. Acad. Sci. USA* 80:7461-7465 (1983)) disclose the use of type-IIs enzymes for the production of vectors to produce recombinant insulin and epidermal growth factor respectively.

[0007] Thus, there remains a need in the art for methods and compositions that allow for insertion of nucleic acid molecules into specific locations of other nucleic acid molecules with seamless junctions at one or both ends. There is also a need in the art for methods and compositions that allow for transfer of these seamlessly cloned sections from one nucleic acid molecule to another. The present invention fulfills these needs.

### Brief Summary of the Invention

[0008] The present invention provides methods of seamlessly cloning nucleic acid molecules. The seamless cloning methods of the present invention may utilize, for example, any restriction enzyme, including those which cleave nucleic acid molecules to produce blunt ends. Suitably, the methods of the invention utilize type IIs restriction sites and enzymes that recognize and cleave at such sites, which allow for the insertion of one or more (e.g. one, two, three, four, five, etc.) nucleic acid segments into specific locations of a



second nucleic acid molecule with seamless junctions on one or both ends. The present methods are also suitable for the production of nucleic acid molecules (e.g. DNA, RNA, DNA hybrids and the like) that only contain nucleic acid sequences that are desired in the product molecule and that lack extraneous unwanted sequences, for example sequences comprising or encoded by restriction sites. The present invention also provides for protein molecules produced or encoded by the cloned nucleic acid molecules of the invention, that contain only amino acid sequences that are desired in the product protein molecule (e.g., a native or mature protein, a fusion protein, and the like), and that lack extraneous amino acids, for example amino acids encoded by restriction sites. In certain embodiments, nucleic acid molecules of the present invention are especially suitable for use as interfering RNA. The present invention also provides novel vectors comprising type II sites and, optionally, selectable markers for the production of seamlessly cloned nucleic acids, as well as compositions and kits for practicing methods of the invention.

[0009] In one aspect, the present invention provides methods for joining one or more (e.g. one, two, three, four, five, etc.) first nucleic acid molecules and one or more second nucleic acid molecules, comprising: (a) combining the first and second nucleic acid molecules under conditions sufficient to allow for the joining of at least one terminus of the first nucleic acid molecule(s) to at least one terminus of the second nucleic acid molecule(s), wherein the terminus of the first nucleic acid molecule(s) which is connected to the terminus of the second nucleic acid molecule(s) comprises a sticky end (e.g. an overhanging end) generated by a restriction enzyme (e.g. a type II restriction enzyme) and the terminus of the second nucleic acid molecule(s) is compatible (e.g. a blunt end or a sticky end) with this sticky end. In embodiments similar to the above and elsewhere herein, the sticky end may be on the terminus of the second nucleic acid molecule, and the first nucleic acid molecule may contain the compatible end.

[0010] In suitable such embodiments, the present invention provides methods of cloning or subcloning one or more desired nucleic acid molecules comprising: (a) combining *in vitro* or *in vivo*, (i) one or more first nucleic acid molecules comprising one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* one or more type II restriction enzymes); and (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more sticky ends on the first nucleic acid molecule(s) and, optionally, one or more selectable markers; and (b) incubating the combination under conditions sufficient to join the first nucleic acid molecule and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0011] In other aspects, the present invention provides methods for cloning or subcloning one or more desired nucleic acid molecules comprising: (a) combining *in vitro* or *in vivo*, (i) one or more first nucleic acid molecules comprising one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* one or more type II restriction enzymes); (ii) one or more second nucleic acid molecules comprising one or more restriction sites (*e.g.* one or more first type II restriction enzyme recognition sites) and, optionally, one or more selectable markers; and (iii) one or more restriction enzymes (*e.g.*, one or more type II restriction enzymes) that are specific for the one or more restriction sites on the second molecules; and (b) incubating the combination under conditions sufficient to join the first nucleic acid molecule and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0012] In additional related aspects, the present invention provides methods for cloning or subcloning one or more desired nucleic acid molecules comprising: (a) combining *in vitro* or *in vivo*, (i) one or more first nucleic acid molecules comprising at least one nucleic acid segment that is flanked by one or more restriction sites (*e.g.* one or more first type II restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with a sticky end on the segment and,

optionally, one or more selectable markers; and (iii) one or more restriction enzymes (*e.g.*, one or more type II restriction enzymes) that are specific for the one or more restriction sites on the at least one nucleic acid segment; and (b) incubating the combination under conditions sufficient to join the first nucleic acid segment and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0013] In related aspects, the present invention provides methods for cloning or subcloning one or more desired nucleic acid molecules, or portions thereof, comprising: (a) combining *in vitro* or *in vivo*, (i) one or more first nucleic acid molecules comprising at least one nucleic acid segment that is flanked by one or more first restriction sites (*e.g.* one or more first type II restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type II restriction enzyme recognition sites) and, optionally, one or more selectable markers; and (iii) one or more restriction enzymes (*e.g.* one or more type II restriction enzymes) that are specific for the first and/or second type II restriction enzyme recognition sites; and (b) incubating the combination under conditions sufficient to join the first nucleic acid segment and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0014] Type II restriction enzyme recognition sites and type II restriction enzymes that are useful in the present cloning methods, compositions, nucleic acids, vectors and kits include, but are not limited to, *Bsa*I, *Bbs*I, *Bbv*II, *Bsm*AI, *Bsp*MI, *Eco*31I, *Bsm*BI, *Bae*I, *Fok*I, *Hga*I, *Mly*I, *Sfa*NI and *Sth*132I. The first, and second restriction sites, if present, utilized throughout the various aspects of the present invention may be the same or they may be different. In addition, the restriction sites on the same nucleic acid molecule (and/or nucleic acid segment) may be the same, or they may be different. The present invention also encompasses situations wherein one or both of the nucleic acid molecules involved in the various methods are vectors, and where one or both of the nucleic acid molecules are linear nucleic acid molecules.

The present invention also encompasses the use of other blunt-end cleavage enzymes, including, but not limited to, *ScaI*, *SmaI*, *HpaI*, *HincII*, *HaeII* and *AluI*.

[0015] In certain embodiments, the nucleic acids and nucleic acid segments utilized in the cloning methods, compositions, kits, and vectors of the present invention may optionally comprise one or more selectable markers. Hence, the invention also provides such nucleic acids. The one or more selectable markers utilized in the present invention may be flanked by one or more (*e.g.* one, two, three, four, five, etc.) restriction sites (*e.g.* type II's restriction enzyme recognition sites). Suitable selectable markers include, but are not limited to, genes that confer antibiotic resistance, genes that encode fluorescent proteins, tRNA genes, auxotrophic markers, toxic genes, phenotypic markers, antisense oligonucleotides, restriction endonucleases, restriction endonuclease cleavage sites, enzyme cleavage sites, protein binding sites, and sequences complementary to PCR primer sequences. Suitable antibiotic resistance genes include, but are not limited to, a chloramphenicol resistance gene, an ampicillin resistance gene, a tetracycline resistance gene, a Zeocin resistance gene, a spectinomycin resistance gene and a kanamycin resistance gene. In certain embodiments of the present invention, the selectable marker is a toxic gene. Suitable toxic genes include, but are not limited to, a *ccdB* gene, a gene encoding a *tus* protein which binds one or more *ter* sites, a *kicB* gene, a *sacB* gene, an ASK1 gene, a  $\Phi$ X174 *E* gene and a DpnI gene. In additional embodiments of the methods of the present invention, the first and/or second nucleic acid molecules may comprise both one or more toxic genes and one or more antibiotic resistance genes, and these genes may further be flanked by type II's restriction enzyme recognition sites. In suitable such embodiments of the present invention, the first and/or second nucleic acid molecules may comprise both a toxic gene and an antibiotic resistance gene.

[0016] In other aspects of the invention, nucleic acids and/or nucleic acid segments for use in the cloning methods, vectors, kits and compositions may

further comprise one or more recombination sites and/or one or more topoisomerase recognition sites and/or one or more topoisomerases. The nucleic acids and/or nucleic acid segments of the present invention may also comprise two or more recombination sites. If a topoisomerase recognition site is present in a nucleic acid molecule or nucleic acid segment of the present invention, it may optionally be flanked by two or more recombination sites. Recombination sites suitable for use in the present invention include, but are not limited to, *attB* sites, *attP* sites, *attL* sites, *attR* sites, *lox* sites, *psi* sites, *tnpI* sites, *dif* sites, *cer* sites, *frt* sites, and mutants, variants and derivatives thereof. These one or more recombination sites may flank one or more selectable markers, if present, and/or restriction sites (e.g. type II sites). In certain embodiments of the present invention, the topoisomerase recognition site, if present, is recognized and bound by a type I topoisomerase, which may be a type IB topoisomerase. Suitable types of type IB topoisomerase include, but are not limited to, eukaryotic nuclear type I topoisomerase and poxvirus topoisomerase. Suitable types of poxvirus topoisomerase include, but are not limited to, poxvirus topoisomerase produced by or isolated from a virus such as vaccinia virus, Shope fibroma virus, ORF virus, fowlpox virus, molluscum contagiosum virus and *Amsacta morrei* entomopoxvirus.

[0017] The present invention also provides methods of linking nucleic acid molecules and/or nucleic acid segments which comprise one or more topoisomerases bound to one or both termini, wherein the topoisomerase adapted terminus or termini comprise a sequence compatible with that cleaved by a restriction enzyme (e.g. a type II restriction enzyme). In such suitable embodiments of the invention, a first nucleic acid molecule or nucleic acid segment may contain a blunt end to be linked, and a second nucleic acid molecule may contain an overhang at the end which is to be linked by a site-specific topoisomerase (e.g., a type IA or a type IB topoisomerase), wherein the overhang includes a sequence complementary to that comprising the blunt end, thereby facilitating strand invasion as a means to properly position the ends for the linking reaction.

[0018] The nucleic acid molecules generated using this aspect of the invention include those in which at least one strand (not both strands) is covalently linked at the ends which are joined (e.g. double-stranded nucleic acid molecules generated using these methods contain a nick at each position where two ends were joined). These embodiments are particularly advantageous in that a polymerase can be used to replicate the double-stranded (ds) nucleic acid molecule by initially replicating the covalently linked strand. For example, a thermostable polymerase such as a polymerase useful for performing an amplification reaction such as PCR can be used to replicate the covalently strand, whereas the strand containing the nick does not provide a suitable template for replication.

[0019] In certain embodiments of the invention, the first or second nucleic acid molecules or nucleic acid segments involved in the various methods of the present invention may not comprise a promoter. The present invention also allows for transfer of a promoter element into a second nucleic acid molecule that may not comprise a promoter, via seamless cloning. In this orientation, transcription of the second nucleic acid molecule from the promoter element located on the first nucleic acid molecule or nucleic acid segment may proceed such that no additional sequences are transcribed between the promoter element and the transcription initiation point of the second nucleic acid molecule. The present invention also allows for seamlessly adding a first nucleic acid molecule or nucleic acid segment into a second nucleic acid molecule that contains a promoter element such that the first nucleic acid molecule or segment will subsequently be under the control of the promoter element.

[0020] The present invention also provides methods for cloning or subcloning one or more desired nucleic acids: (a) combining *in vitro* or *in vivo*, (i) one or more first nucleic acid molecules that have one or more sticky ends that have been generated by one or more restriction enzymes (e.g. type II restriction enzymes); and (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more sticky ends on the

first nucleic acid molecule(s) and further comprising one or more recombination sites; and (b) incubating the combination under conditions sufficient to join the first nucleic acid molecule and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0021] The present invention also provides methods for cloning or subcloning one or more desired nucleic acid molecules, or portions thereof, comprising: (a) combining *in vitro* or *in vivo*, (i) one or more first nucleic acid molecules comprising at least one nucleic acid segment that is flanked by one or more first restriction sites (*e.g.* one or more type IIs restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* type IIs restriction enzyme recognition sites) flanked by one or more recombination sites; and (iii) one or more restriction enzymes (*e.g.* one or more type IIs restriction enzymes) that are specific for the first and/or second restriction sites; and (b) incubating the combination under conditions sufficient to join the first nucleic acid molecule and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0022] As described above, the first and/or second nucleic acid molecules and/or nucleic acid segments involved in such embodiments of the present invention may optionally comprise one or more selectable markers. The first and/or second nucleic acid molecules and/or nucleic acid segments involved in such aspects of the invention may also, or alternatively comprise one or more topoisomerase recognition sites or topoisomerases as described above, and optionally or alternatively, two or more recombination sites, which in certain such embodiments may flank these topoisomerases or topoisomerase recognition sites.

[0023] The present invention also provides methods for cloning or subcloning one or more desired nucleic acid molecules, or portions thereof, via recombination cloning comprising: (a) combining, *in vitro* or *in vivo* (i) one or more first nucleic acid molecules comprising at least one nucleic acid segment

that is flanked by one or more restriction sites (*e.g.* one or more type II<sub>s</sub> restriction enzyme recognition sites) and that is further flanked by one or more recombination sites; (ii) one or more second nucleic acid molecules comprising one or more recombination sites; and (iii) one or more site-specific recombination proteins; and (b) incubating the combination under conditions sufficient to join the first nucleic acid molecule and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0024] The second nucleic acid molecule involved in such embodiments of the invention may also comprise one or more restriction sites (*e.g.* one or more type II<sub>s</sub> restriction enzyme recognition sites). The first and/or second nucleic acids and/or nucleic acid segments involved may also optionally comprise one or more selectable markers as described above. The first and/or second nucleic acid molecules and/or nucleic acid segments involved in this aspect of the invention may also comprise topoisomerase recognition sites or topoisomerases as described above, as well as two or more recombination sites flanking these topoisomerase sites.

[0025] Suitable recombination proteins for use in the present invention include, but are not limited to, Int, Cre, IHF, Xis, Fis, Hin, Gin, Cin, Tn3 resolvase, TndX, XerC and XerD.

[0026] The present invention also provides methods for producing host cells comprising one or more of the nucleic acid molecules produced by the cloning methods of the present invention. Suitable host cells that may be used throughout the present invention include, but are not limited to, bacterial cells, yeast cells, plant cells and animal cells. The present invention also provides methods for producing a subsequent nucleic acid molecule and/or protein by expression of the product nucleic acid molecule of the cloning methods of the present invention in a host cell.

[0027] Additional embodiments provide for nucleic acid molecules and proteins produced in and isolated from a host cell. In certain such embodiments, the nucleic acid molecules produced in the host cell may



contain only desired nucleic acid sequences, *i.e.* they may not contain extraneous nucleotides, for example, nucleotides encoded by the restriction sites (*e.g.* type II restriction enzyme recognition sites). Similarly, the proteins produced from a host cell by these methods may only contain amino acid sequences that correspond to the desired native or mature protein, and may not contain extraneous amino acids, for example amino acids encoded by the restriction sites (*e.g.* type II restriction enzyme recognition sites). Nucleic acid molecules produced from a host cell by methods of the present invention may be useful as interfering RNA molecules.

[0028] Another aspect of the present invention provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules encoding one or more interfering RNAs that have one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* type II restriction enzymes); and (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more sticky ends on the first nucleic acid molecule(s), and optionally comprising one or more selectable markers; and (d) incubating the combination under conditions sufficient to join one or more of the nucleic acid molecules encoding the interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; (e) inserting the one or more product nucleic acid molecules into a host cell; and (f) expressing the one or more interfering RNAs in the host cell. As in other embodiments of the invention described herein, the second nucleic acid molecule may contain an end which is generated by digestion with a type II restriction enzyme and the first nucleic acid molecule may contain a compatible end generated by other means.

[0029] The present invention also provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules encoding one or more interfering RNAs flanked by one or more first restriction sites (*e.g.* one or more type IIs restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type IIs restriction enzyme recognition sites) and optionally comprising one or more selectable markers; and (iii) one or more site-specific restriction enzymes (*e.g.* one or more type IIs restriction enzymes); and (d) incubating the combination under conditions sufficient to join one or more of the nucleic acid molecules encoding the interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; (e) inserting the one or more product nucleic acid molecules into a host cell; and (f) expressing the one or more interfering RNAs in the host cell.

[0030] In related embodiments, the present invention provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules encoding one or more interfering RNAs that have one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* type IIs restriction enzymes); and (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more sticky ends on the first nucleic acid molecule(s), and optionally comprising one or more selectable markers; and (d) incubating the combination under conditions sufficient to join one or more of the nucleic acid

molecules encoding the interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; and (e) expressing one or more interfering RNAs *in vitro* or *in vivo*. In a first further embodiment, the one or more interfering RNAs may be produced *in vitro* or isolated from a cell and then introduced into a second cell.

[0031] Another aspect of the present invention provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules encoding one or more interfering RNAs flanked by one or more first restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type II restriction enzyme recognition sites) and optionally comprising one or more selectable markers; and (iii) one or more site-specific restriction enzymes (*e.g.* one or more type II restriction enzymes); and (d) incubating the combination under conditions sufficient to join one or more of the nucleic acid molecules encoding the interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; and (e) expressing one or more interfering RNAs *in vitro* or *in vivo*. In a first further embodiment, the one or more interfering RNAs may be produced *in vitro* or isolated from a cell and then introduced into a second cell.

[0032] In a related aspect, the present invention provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in*

*vivo*, (i) the one or more first nucleic acid molecules comprising one or more interfering RNAs that have one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* type II restriction enzymes); and (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more sticky ends on the first nucleic acid molecule(s), and optionally comprising one or more selectable markers; and (d) incubating the combination under conditions sufficient to join one or more interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; (e) inserting the one or more product nucleic acid molecules into a host cell; and (f) expressing the one or more interfering RNAs in the host cell.

[0033] The present invention also provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which comprise one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules comprising one or more interfering RNAs flanked by one or more first restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type II restriction enzyme recognition sites) and optionally comprising one or more selectable markers; and (iii) one or more site-specific restriction enzymes (*e.g.* one or more type II restriction enzymes); and (d) incubating the combination under conditions sufficient to join one or more interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; (e) inserting the one or more product nucleic acid molecules into a host cell; and (f) expressing the one or more interfering RNAs in the host cell.

[0034] Methods of the present invention may be used, for example, to prepare shRNA molecules in which the 5' and 3' termini contain none or few (*e.g.*, one, two, three, four, or five) nucleotides which are not encoded by a first

nucleic acid molecule referred to throughout. Thus, the shRNA may comprise from about 40 to about 60 nucleotides in which either none of all but a few nucleotides at one or both termini are encoded by a first nucleic acid molecule. In such instances, the first nucleic acid molecule may be composed of nucleic acid which upon transcription results in the production of RNA with three different segments: (1) sense RNA, (2) a loop/non-complementary RNA, and (3) antisense RNA. Methods of the invention include introducing into a cell (1) (a) nucleic acid which encodes the RNA described above or (b) the RNA itself, and (2) the measurement of inhibition of expression of a gene corresponding to the sense and/or antisense RNA.

[0035] In particular embodiments of the invention, the invention may be used to produce nucleic acid molecules which produce RNA molecules that do not form hairpins. As one example, methods of the invention may be used to produce two separate vectors, one or which may be used to produce a sense RNA molecules (*e.g.*, a sense RNA molecule which is between about 18 and about 30, between about 20 and about 30, between about 22 and about 30, or between about 18 and about 25 nucleotides in length) and an antisense RNA molecules (*e.g.*, a sense RNA molecule which is between about 18 and about 30, between about 20 and about 30, between about 22 and about 30, between about 18 and about 100, or between about 18 and about 25 nucleotides in length), wherein the two RNA molecules are capable of hybridizing to each other and/or share a region of sequence complementarity over at least 80%, 90%, or 95% of their full lengths (*e.g.*, sequence complementarity over a 19 nucleotide stretch, wherein each molecule is 22 nucleotides in length). Alternatively, both sense and antisense RNA molecules, such as described above, may be produced by a single vector but as separate transcription products.

[0036] As a variation of the above, the invention may be used to produce either sense or antisense RNA molecules alone in cells. These RNA molecules may be of any length suitable for the particular application (*e.g.*,

expression of protein, antisense inhibition of gene expression, ribozyme production, etc.).

[0037] The invention may further be used to produce microRNA molecules. MicroRNA molecules are molecules which are structurally similar to shRNA molecules but, typically, contain one or more mismatches or insertion/deletions in their regions of sequence complementary. At least some microRNA molecules are transcribed as polycistrons of about 400, which are then processed to RNA molecules of about 70 nucleotides. These double stranded 70 mers are then are processed again, presumably by the enzyme Dicer, to two RNA molecules which are about 22 nucleotides in length and often have one or more (*e.g.*, one, two, three, four, five, etc.) internal mismatches in their regions of sequence complementarity. Lee *et al.*, *EMBO* 21:4663-4670 (2002). The invention also includes, for example, uses of microRNA molecules and nucleic acid molecules which encode microRNA molecules which are similar to the uses described those described herein for shRNA and non-hairpin double stranded RNA molecules.

[0038] The present invention also provides methods of regulating the expression of one or more genes in a cell or an animal using interfering RNA, comprising: (a) identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules encoding one or more interfering RNAs that have one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* type II restriction enzymes); and (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more sticky ends on the first nucleic acid molecule(s), and optionally comprising one or more selectable markers; (d) incubating the combination under conditions sufficient to join one or more of the nucleic acid molecules encoding the interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic

acid molecules; and (e) inserting the one or more interfering RNA expression vectors into the cell or one or more cells of the animal, under conditions such that the one or more interfering RNAs bind to the one or more target nucleic acid sequences, thereby regulating expression of the one or more targeted genes.

[0039] In related embodiments, the present invention also provides methods of regulating the expression of one or more genes in a cell or an animal using interfering RNA, comprising: (a) identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which comprise one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules comprising one or more interfering RNAs flanked by one or more first restriction sites (*e.g.* one or more type IIs restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type IIs restriction enzyme recognition sites) and optionally comprising one or more selectable markers; and (iii) one or more site-specific restriction enzymes (*e.g.* one or more type IIs restriction enzymes); (d) incubating the combination under conditions sufficient to join one or more interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; and (e) inserting the one or more interfering RNA expression vectors into the cell or one or more cells of the animal, under conditions such that the one or more interfering RNAs bind to the one or more target nucleic acid sequences, thereby regulating expression of the one or more targeted genes.

[0040] Such methods of the invention can be used to knockout or knockdown one or more genes *in vivo* in a cell or animal. These methods of the invention may also be used to produce genetically modified animals by expressing interfering RNA in germ cells or somatic cells, and for preparation of transgenic animals.

[0041] In another embodiment, the present invention also provides isolated nucleic acid molecules comprising: (a) one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* one or more type II restriction enzymes); and (b) optionally one or more selectable markers. The present invention also provides isolated nucleic acid molecules comprising: (a) one or more restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); and (b) optionally one or more selectable markers.

[0042] Suitable restriction enzyme recognition sites and selectable markers are described above. The isolated nucleic acid molecules of the present invention may also comprise one or more recombination sites and/or one or more topoisomerase recognition sites and/or one or more topoisomerases. If present, the topoisomerase recognition sites may be flanked by recombination sites. The isolated nucleic acid molecules of the present invention may be vectors or linear nucleic acid molecules. The present invention also provides isolated nucleic acid molecules comprising: (a) one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* one or more type II restriction enzymes); and (b) one or more recombination sites. The present invention further provides isolated nucleic acid molecules comprising: (a) one or more restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); and (b) one or more recombination sites.

[0043] The present invention also provides vectors comprising: (a) one or more desired nucleic acid segments; (b) optionally one or more toxic genes; and (c) one or more sites that are compatible with a sticky end generated by a restriction enzyme (*e.g.* one or more type II restriction enzymes). Suitable desired nucleic acid molecules include genes (*e.g.* open reading frames) and promoters. The vectors of the present invention may also comprise one or more recombination sites, and one or more topoisomerase recognition sites and/or one or more topoisomerases, wherein, the topoisomerase recognition sites if present, may be flanked by recombination sites. In other embodiments, the vectors of the present invention may optionally comprise one or more selectable markers as described above. Suitable vectors of the present



invention include, but are not limited to, pENTR/U6-*ccdB* (vector diagram shown in Figure 2A, vector sequence in Table 5 and SEQ ID NO:1).

[0044] The present invention also provides vectors comprising: (a) one or more desired nucleic acid segments; (b) optionally one or more toxic genes; and (c) one or more restriction sites (*e.g.* one or more type II restriction enzyme recognition sites). Suitable desired nucleic acid molecules include genes and promoters. The vectors of the present invention may also comprise one or more recombination sites, and one or more topoisomerase recognition sites and/or one or more topoisomerases, wherein, the topoisomerase recognition sites if present, may be flanked by recombination sites. In other embodiments, the vectors of the present invention may optionally comprise one or more selectable markers as described above. Suitable vectors of the present invention include, but are not limited to, pENTR/U6-*ccdB* (vector diagram shown in Figure 2A, vector sequence in Table 5, Figure 12 and SEQ ID NO:1).

[0045] The present invention also provides host cells comprising one or more of the isolated nucleic acid molecules or nucleic acid segments of the present invention, and methods of expressing the isolated nucleic acids of the present invention in one more host cells and isolating the expressed nucleic acids. The present invention also provides methods of expressing and isolating proteins from host cells comprising one or more isolated nucleic acids or nucleic acid segments of the invention.

[0046] Another embodiment of the invention provides methods of expressing desired product nucleic acid segments by introducing the nucleic acid molecules, nucleic acid segments, or vectors of the present invention into a host cell and expressing the product nucleic acid segments.

[0047] The present invention also provides for compositions comprising: (a) one or more first nucleic acid molecules that have one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* type II restriction enzymes); and (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more

sticky ends on the first nucleic acid molecule(s). The first and second nucleic acid molecules may optionally comprise one or more selectable markers as discussed above. These first and second nucleic acid molecules may also comprise one or more recombination sites, one or more topoisomerase recognition sites and/or one or topoisomerases, wherein the topoisomerase recognition sites, if present, may be flanked by recombination sites. The optional selectable markers may be flanked by type II restriction sites and/or recombination sites. The compositions of the invention may also comprise one or more recombination proteins as described above.

[0048] The present invention further provides for compositions comprising: (a) one or more first nucleic acid molecules comprising at least one nucleic acid segment that is flanked by one or more first restriction sites (e.g. one or more type II restriction enzyme recognition sites; (b) one or more second nucleic acid molecules optionally comprising one or more second restriction sites (e.g. one or more type II restriction enzyme recognition sites); and (c) one or more restriction enzymes (e.g. type II restriction enzymes) that are specific for the first and/or second restriction sites. The first and second nucleic acid molecules and/or nucleic acid segments may optionally comprise one or more selectable markers as discussed above. These first and second nucleic acid molecules and/or nucleic acid segments may also comprise one or more recombination sites, one or more topoisomerase recognition sites and/or one or topoisomerases, wherein the topoisomerase recognition sites, if present, may be flanked by recombination sites. The optional selectable markers may be flanked by type II restriction sites and/or recombination sites. The compositions of the invention may also comprise one or more recombination proteins as described above.

[0049] The present invention also provides kits comprising the isolated nucleic acids or vectors of the present invention. The kits of the present invention may further comprise one or more type II restriction enzymes, one or more recombination proteins, and one or more host cells.

[0050] Other embodiments of the present invention will be apparent to one or ordinary skill in light of the following drawings and description of the invention, and of the claims.

### Brief Description of the Drawings

[0051] **Figure 1A** is a schematic diagram of a vector of the invention comprising: an origin of replication (ori), a kanamycin resistance gen (kan), a Polymerase II promoter (polII), L1 (*attL1*) and L2 (*attL2*) recombination sites, an ATG translation initiation site/codon, a secretion signal, type IIs restriction sites, and a negative selectable marker.

[0052] **Figure 1B** is a schematic diagram of a vector of the invention comprising: an origin of replication (ori), a kanamycin resistance gen (kan), a Polymerase II promoter (polII), L1 (*attL1*) and L2 (*attL2*) recombination sites, an ATG initiation site/codon, an affinity tag, a cleavage site, a type IIs restriction site, and a negative selectable marker.

[0053] **Figure 2A** is a schematic diagram of pENTR/U6.

[0054] **Figure 2B** depicts a *BsaI* digestion and cloning scheme using pENTR/U6.

[0055] **Figures 3A and 3B** depict luciferase and  $\beta$ -gal suppression in Griptite™ 293 cells by transient cotransfection of reporters and pENTR/U6 vectors. A) Luciferase activities measured in lysates of cells: from left 1) untransfected, 2) cotransfected with luciferase and lacZ reporter genes plus a dummy plasmid (pUC19/actin), or 3-4) same as 2 except either pENTR/U6 targeting luciferase (GL2-22) or  $\beta$ -gal (lacZ-19) replace the pUC19/actin. B)  $\beta$ -gal activity measurements of the same lysates as in A. Activities are the average of duplicate wells. The standard error of the mean is indicated for each sample.

[0056] **Figure 4A and 4B** depicts RNAi of  $\beta$ -Gal and Luciferase activity from co-transfected reporter constructs by pENTR/U6 shRNA clones. Data are reported as the ratio of lacZ and Luciferase activity. Error bars are calculated

from two independent samples. AS/SA indicates the orientation of the sense and anti-sense strand relative to the U6 promoter. **A)** Luciferase/ $\beta$ -gal activity after co-transfection with the indicated pENTR/U6 shRNA sequences targeting the Luciferase gene and a pUC19-actin control. pENTR/U6-A6-GL2-22(AS) is the same construct used in **Figure 3**. The asterisk (\*) after ENT/6-A6-GL2-2-SA indicates a point mutation was identified in the shRNA target sequence clone used in this experiment. **B)**  $\beta$ -gal/Luciferase activity after co-transfection with various pENTR/U6 shRNA sequences targeting the LacZ gene. ENT/6-A6-lacZ-19 is the same construct used to generate the data presented in **Figure 3**.

[0057] **Figure 5** depicts  $\beta$ -gal/Luciferase activity ratios after co-transfection reporter plasmids and pENTR/U6 LacZ-19 shRNA target clones with the indicated Terminator lengths. Terminators with 4, 5, 6 and 8 "Ts" were tested in the pENTR/U6.2 vector (A4-8).

[0058] **Figure 6A** is a schematic of the lentiviral RNAi shRNA transfer vector: pLenti6/RNAi-DEST which is a promoterless Gateway-adapted lenti vector which may be used to clone, for example an shRNA cassette of interest via Gateway LxR reaction with pENTR U6 vectors. The shRNA cassette will often contain an RNA pol III - or other- promoter of choice to drive hairpin expression. The vector confers blasticidin resistance to transduced cells.

[0059] **Figure 6B** is a schematic of the lentiviral RNAi Kit control vector: Kit control plasmid pLenti6/RNAi/U6-GW/lamAC which results from LxR reaction between pLenti6/RNAi-DEST and pENTR/U6-lamAC-AS-cgaa. pLenti6/RNAi/U6-GW/lamAC expresses lamAC-AS-cgaa hairpin to specifically knockdown lamin A/C expression.

[0060] **Figure 7** depicts the inhibition of lamin A/C expression. Lenti6/RNAi viruses encoding anti-lamin A/C shRNAs (U6-lamAC) were transduced into HeLa cells to test inhibition of lamin A/C expression. Control viruses encoded GFP gene (GFP) or anti-luciferase shRNAs (U6-GL2). Western blots for lamin A/C or beta-actin were conducted on lysates from transduced cells. Top panel: Lysates were prepared 48 hrs post-transduction. Bottom panel:

Lysates were prepared from transduced, shRNA-producing, blasticidin-resistant cells 5 days post-transduction.

- [0061]      **Figure 8A** is a plasmid map of pLenti6/V5-DEST.
- [0062]      **Figure 8B** is a plasmid map of pLenti6/V5-gTOPO®.
- [0063]      **Figure 8C** is a plasmid map of pLenti4/V5-DEST
- [0064]      **Figure 8D** is a plasmid map of pLenti6/UbC/V5-DEST.
- [0065]      **Figure 9A** is a plasmid map pLP1.
- [0066]      **Figure 9B** is a plasmid pLP2.
- [0067]      **Figure 9C** is a plasmid map of pLP/VSVG.
- [0068]      **Figure 10** is a plasmid map of pAd/PL-DEST.
- [0069]      **Figure 11** is a plasmid map of pAd/CMV/V5-DEST.
- [0070]      **Figure 12** depicts the nucleic acid sequence of the pENTR/U6 with annotations noting the various segments of the vector. SEQ ID NO:1
- [0071]      **Figure 13** depicts RNAi overview.
- [0072]      **Figure 14** depicts RNAi Mechanistic Model.
- [0073]      **Figure 15** depicts RNAi Methods.
- [0074]      **Figure 16** depicts siRNA Molecules.
- [0075]      **Figure 17** depicts Transfection of siRNAs
- [0076]      **Figure 18** depicts Variation in siRNA effectiveness.
- [0077]      **Figure 19** depicts expression *in vivo*.
- [0078]      **Figure 20** depicts BLOCK-iT™ Long RNAi Transcription Kit.
- [0079]      **Figure 21** depicts BLOCK-iT™ Dicer RNAi Kit
- [0080]      **Figure 22** depicts d-siRNA knockdown.
- [0081]      **Figure 23** depicts d-siRNA vs. siRNA.
- [0082]      **Figure 24** depicts BLOCK-iT™ RNAi.
- [0083]      **Figure 25** depicts Micro RNA (miRNA).
- [0084]      **Figure 26** depicts RNAi Vectors.
- [0085]      **Figure 27** depicts U6 RNAi.
- [0086]      **Figure 28** depicts Gateway™ Cloning and ViraPower™ RNAi cassettes.
- [0087]      **Figure 29** depicts Selecting a viral expression system.

- [0088]        **Figure 30** depicts Outline for lentiviral production.
- [0089]        **Figure 31** depicts Overview of Lentiviral Production.
- [0090]        **Figure 32** depicts ViraPower™ lentiviral production.
- [0091]        **Figure 33** depicts Clone your gene of interest into Lentivirus.
- [0092]        **Figure 34** depicts Two methods for fast cloning.
- [0093]        **Figure 35** depicts Two methods for fast cloning.
- [0094]        **Figure 36** depicts Subcloning an Entry Clone into Multiple  
Destination Vectors.
- [0095]        **Figure 37** depicts pLenti6/V5 Expression Vectors.
- [0096]        **Figure 38** depicts GATEWAY Cloning Technology.
- [0097]        **Figure 39** depicts Assembly of Three DNA segments using Existing  
Entry Clones.

## Detailed Description of the Invention

### Definitions

- [0098]        Unless defined otherwise, all technical and scientific terms used herein  
have the same meanings as commonly understood by one of ordinary skill in  
the art to which this invention belongs.
- [0099]        One or more: As used herein, the term “one or more” includes at least  
one, more suitably, one, two, three, four, five, ten, twenty, fifty, one-hundred,  
five-hundred, etc., of the item to which “one or more” refers.
- [0100]        Nucleic Acid: As used herein, “nucleic acid” refers to polynucleotides  
such as deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). The term  
should also be understood to include, as applicable to the embodiment being  
described, single-stranded (such as sense or antisense) and double stranded  
polynucleotides, including double-stranded DNA-RNA hybrids. The term  
“nucleic acid” also is synonymous, and may be used interchangeably with the  
term “nucleic acid molecule.”

- [0101] Gene: As used herein, "gene" refers to a nucleic acid comprising an open reading frame encoding a polypeptide, including both exon and (optionally) intron sequences.
- [0102] About: As used herein, when referring to any numerical value, "about" means a value of  $\pm 10\%$  of the stated value (e.g. "about 50°C encompasses a range of temperatures from 45°C to 55°C, inclusive: similarly, "about 100 mM" encompasses a range of concentrations from 90 mM to 110 mM, inclusive).
- [0103] Host: As used herein, a "host" is any prokaryotic or eukaryotic organism that is a recipient of a replicable expression vector, cloning vector or any nucleic acid molecule. The nucleic acid molecule may contain, but is not limited to, a structural gene, a transcriptional regulatory sequence (such as a promoter, enhancer, repressor, and the like) and/or an origin of replication. As used herein, the terms, "host," "host cell," "recombinant host" and "recombinant host cell" may be used equivalently and interchangeably. For examples of such hosts, see Maniatis et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982).
- [0104] Derivative: As used herein the term "derivative," when used in reference to a vector, means that the derivative vector contains one or more (e.g., one, two, three, four five, etc.) nucleic acid segments which share sequence similar to the vectors represented in Figure 1A, Figure 1B, Figure 2A, Figure 6A, Figure 6B, Figure 8A, Figure 8B, Figure 8C, Figure 8D, Figure 9A, Figure 9B, Figure 9C, Figure 10, Figure 11, Figure 12, Table 5, and any other vector encompassed by the present application. In particular embodiments, a derivative vector (1) may be obtained by alteration of a vector represented in Figure 1A, Figure 1B, Figure 2A, Figure 6A, Figure 6B, Figure 8A, Figure 8B, Figure 8C, Figure 8D, Figure 9A, Figure 9B, Figure 9C, Figure 10, Figure 11, Figure 12, Table 5, and any other vector encompassed by the present application, or (2) may contain one or more elements (e.g., antibiotic resistance marker, recombination or restriction site,

etc.) of a vector represented in Figure 1A, Figure 1B, Figure 2A, Figure 6A, Figure 6B, Figure 8A, Figure 8B, Figure 8C, Figure 8D, Figure 9A, Figure 9B, Figure 9C, Figure 10, Figure 11, Figure 12, Table 5, and any other vector encompassed by the present application. Further, as noted above, a derivative vector may contain one or more element which shares sequence similarity (e.g., at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, etc. sequence identity at the nucleotide level) to one or more element of a vector represented in Figure 1A, Figure 1B, Figure 2A, Figure 6A, Figure 6B, Figure 8A, Figure 8B, Figure 8C, Figure 8D, Figure 9A, Figure 9B, Figure 9C, Figure 10, Figure 11, Figure 12, Table 5, and any other vector encompassed by the present application. Derivative vectors may also share at least at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, etc. sequence identity at the nucleotide level to the complete nucleotide sequence of a vector represented in Figure 1A, Figure 1B, Figure 2A, Figure 6A, Figure 6B, Figure 8A, Figure 8B, Figure 8C, Figure 8D, Figure 9A, Figure 9B, Figure 9C, Figure 10, Figure 11, Figure 12, Table 5, and any other vector encompassed by the present application. Derivative vectors include those which have been generated by performing a cloning reaction upon a vector represented in Figure 1A, Figure 1B, Figure 2A, Figure 6A, Figure 6B, Figure 8A, Figure 8B, Figure 8C, Figure 8D, Figure 9A, Figure 9B, Figure 9C, Figure 10, Figure 11, Figure 12, Table 5, and any other vector encompassed by the present application. Derivative vectors also include vectors which have been generated by the insertion into another vector of one or more structural and/or functional components of a vector (e.g. one or more genes or portions thereof encoding one or more structural or functional proteins (or portions thereof) of a vector), including but not limited to the vectors represented in Figure 1A, Figure 1B, Figure 2A, Figure 6A, Figure 6B, Figure 8A, Figure 8B, Figure 8C, Figure 8D, Figure 9A, Figure 9B, Figure 9C, Figure 10, Figure 11, Figure 12, Table 5, and any other vector encompassed by or suitable for use in the invention. Often these derivative vectors will contain at least 50%, at least



60%, at least 70%, at least 80%, at least 90%, at least 95%, etc. of the nucleic acid present in a vector represented in Figure 1A, Figure 1B, Figure 2A, Figure 6A, Figure 6B, Figure 8A, Figure 8B, Figure 8C, Figure 8D, Figure 9A, Figure 9B, Figure 9C, Figure 10, Figure 11, Figure 12, Table 5, and any other vector encompassed by the present application. Derivative vectors also include progeny of any of the vectors referred to above, as well as vectors referred to above which have been subjected to mutagenesis (e.g., random mutagenesis).

[0105] Promoter: As used herein, a promoter is an example of a transcriptional regulatory sequence, and is specifically a nucleic acid sequence generally described as the proximal region of a gene located 5' to the start codon. The transcription of an adjacent nucleic acid segment is initiated at the promoter region. A repressible promoter's rate of transcription decreases in response to a repressing agent. An inducible promoter's rate of transcription increases in response to an inducing agent. A constitutive promoter's rate of transcription is not specifically regulated, though it can vary under the influence of general metabolic conditions. Suitable examples of promoters that may be used in the present invention include, but are not limited to polymerase III promoters such as H1 and U6.

[0106] Product: As used herein, a "product" is one of the desired daughter molecules produced after cloning process. The product contains the nucleic acid which was to be cloned or subcloned.

[0107] Recognition sequence: As used herein, a "recognition sequence" (alternatively and equivalently referred to herein as a recognition site) is a particular sequence to which a protein, chemical compound, DNA, or RNA molecule (e.g., restriction endonuclease, a topoisomerase, a modification methylase, a type IIs restriction enzyme, or a recombinase) recognizes and binds. In the present invention, a recognition sequence may refer to a recombination site (which may alternatively be referred to as a recombinase recognition site), a topoisomerase recognition site, or a type IIs restriction enzyme recognition site. For example, the recognition sequence for Cre

recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Current Opinion in Biotechnology* 5:521-527 (1994). Other examples of such recognition sequences are the *attB*, *attP*, *attL*, and *attR* sequences which are recognized by the recombinase enzyme. Integrase *attB* is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region. *attP* is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). See Landy, *Current Opinion in Biotechnology* 3:699-707 (1993). Such sites may also be engineered according to the present invention to enhance production of products in the methods of the invention. When such engineered sites lack the P1 or H1 domains to make the recombination reactions irreversible (e.g., *attR* or *attP*), such sites may be designated *attR'* or *attP'* to show that the domains of these sites have been modified in some way. Examples of topoisomerase recognition sites include, but are not limited to, the sequence 5'-GCAACTT-3' that is recognized by *E. coli* topoisomerase III (a type I topoisomerase); the sequence 5'-(C/T)CCTT-3' which is a topoisomerase recognition site that is bound specifically by most poxvirus topoisomerases, including vaccinia virus DNA topoisomerase I; and others that are known in the art as discussed elsewhere herein.

[0108]       Recombination proteins: As used herein, "recombination proteins" include excisive or integrative proteins, enzymes, co-factors or associated proteins that are involved in recombination reactions involving one or more recombination sites, which may be wild-type proteins (See Landy, *Current Opinion in Biotechnology* 3:699-707 (1993)), or mutants, derivatives (e.g., fusion proteins containing the recombination protein sequences or fragments thereof), fragments, and variants thereof. Suitable recombination proteins for use in the present invention include, but are not limited to Int, Cre, IHF, Xis, Fis, Hin, Gin, Cin, Tn3 resolvase, TndX, XerC and XerD.

- [0109]        Recombination site: As used herein, a “recombination site” is a recognition sequence on a nucleic acid molecule participating in an integration/recombination reaction by recombination proteins. Recombination sites are discrete sections or segments of nucleic acid on the participating nucleic acid molecules that are recognized and bound by a site-specific recombination protein during the initial stages of integration or recombination. For example, the recombination site for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Curr. Opin. Biotech.* 5:521-527 (1994). Other examples of recognition sequences include the *attB*, *attP*, *attL*, and *attR* sequences described herein, and mutants, fragments, variants and derivatives thereof, which are recognized by the recombination protein *Int* and by the auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). See Landy, *Curr. Opin. Biotech.* 3:699-707 (1993).
- [0110]        Recombinational Cloning: As used herein, “recombinational cloning” is a method, such as that described in U.S. Patent Nos. 5,888,732, 6,143,557, 6,171,861, 6,270,969, and 6,277,608 (the contents of which are fully incorporated herein by reference), whereby segments of nucleic acid molecules or populations of such molecules are exchanged, inserted, replaced, substituted or modified, *in vitro* or *in vivo*. Suitably, such cloning method is an *in vitro* method, *i.e.*, a method in which the recombination reaction takes place outside of or in the absence of host cells.
- [0111]        Selectable marker: As used herein, “selectable marker” is a nucleic acid segment that allows one to select for or against a molecule (e.g., a replicon) or a cell that contains it, often under particular conditions. These markers can encode an activity, such as, but not limited to, production of RNA, peptide, or protein, or can provide a binding site for RNA, peptides, proteins, inorganic and organic compounds or compositions and the like. Examples of selectable markers include but are not limited to: (1) nucleic acid segments that encode products which provide resistance against otherwise

toxic compounds (e.g., antibiotics); (2) nucleic acid segments that encode products which are otherwise lacking in the recipient cell (e.g., tRNA genes, auxotrophic markers); (3) nucleic acid segments that encode products which suppress the activity of a gene product; (4) nucleic acid segments that encode products which can be readily identified (e.g., phenotypic markers such as  $\beta$ -galactosidase, green fluorescent protein (GFP), yellow fluorescent protein (YFP), cyan fluorescent protein (CFP), and cell surface proteins); (5) nucleic acid segments that bind products which are otherwise detrimental to cell survival and/or function; (6) nucleic acid segments that otherwise inhibit the activity of any of the nucleic acid segments described in Nos. 1-5 above (e.g., antisense oligonucleotides); (7) nucleic acid segments that bind products that modify a substrate (e.g. restriction endonucleases); (8) nucleic acid segments that can be used to isolate or identify a desired molecule (e.g. specific protein binding sites); (9) nucleic acid segments that encode a specific nucleotide sequence which can be otherwise non-functional (e.g., for PCR amplification of subpopulations of molecules); (10) nucleic acid segments, which when absent, directly or indirectly confer resistance or sensitivity to particular compounds; and/or (11) nucleic acid segments that encode products which are toxic in recipient cells.

[0112] Examples of toxic gene products are well known in the art, and include, but are not limited to, restriction endonucleases (e.g., DpnI), apoptosis-related genes (e.g. ASK1 or members of the bcl-2/ced-9 family), retroviral genes including those of the human immunodeficiency virus (HIV), defensins such as NP-1, inverted repeats or paired palindromic nucleic acid sequences, bacteriophage lytic genes such as those from ( $\Phi$ X174 or bacteriophage T4; antibiotic sensitivity genes such as rpsL, antimicrobial sensitivity genes such as pheS, plasmid killer genes, eukaryotic transcriptional vector genes that produce a gene product toxic to bacteria, such as GATA-1, and genes that kill hosts in the absence of a suppressing function, e.g., *kicB*, *ccdB*,  $\Phi$ X174 *E* (Liu, Q. *et al.*, *Curr. Biol.* 8:1300-1309 (1998), and other

genes that negatively affect replicon stability and/or replication. A toxic gene can alternatively be selectable *in vitro*, e.g., a restriction site.

[0113] Selection scheme: As used herein, "selection scheme" is any method which allows selection, enrichment, or identification of a desired product or product(s). The selection schemes of one suitable embodiment have at least two components that are either linked or unlinked during recombinational cloning. One component is a Selectable marker. The other component controls the expression *in vitro* or *in vivo* of the Selectable marker, or survival of the cell (or the nucleic acid molecule, e.g., a replicon) harboring the plasmid carrying the Selectable marker. Generally, this controlling element will be a repressor or inducer of the Selectable marker, but other means for controlling expression or activity of the Selectable marker can be used. Whether a repressor or activator is used will depend on whether the marker is for a positive or negative selection, and the exact arrangement of the various nucleic acid segments, as will be readily apparent to those skilled in the art.

[0114] Fragments of selectable markers can be arranged relative to the recombination sites or restriction sites such that when the segments are brought together, they reconstitute a functional Selectable marker. For example, the linking event can link a promoter with a structural nucleic acid molecule (e.g., a gene), can link two fragments of a structural nucleic acid molecule, or can link nucleic acid molecules that encode a heterodimeric gene product needed for survival, or can link portions of a replicon.

[0115] Site-specific recombinase: As used herein, a "site specific recombinase" is a type of recombinase which typically has at least the following four activities (or combinations thereof): (1) recognition of one or two specific nucleic acid sequences; (2) cleavage of said sequence or sequences; (3) topoisomerase activity involved in strand exchange; and (4) ligase activity to reseal the cleaved strands of nucleic acid. See Sauer, B., *Current Opinions in Biotechnology* 5:521-527 (1994). Conservative site-specific recombination is distinguished from homologous recombination and transposition by a high degree of specificity for both partners. The strand

exchange mechanism involves the cleavage and rejoining of specific nucleic acid sequences in the absence of DNA synthesis (Landy, A. (1989) *Ann. Rev. Biochem.* 58:913-949).

[0116] Vector: As used herein, a “vector” is a nucleic acid molecule (preferably DNA) that provides a useful biological or biochemical property to an Insert. Examples include plasmids, phages, autonomously replicating sequences (ARS), centromeres, and other sequences which are able to replicate or be replicated *in vitro* or in a host cell, or to convey a desired nucleic acid segment to a desired location within a host cell. A Vector can have one or more restriction endonuclease recognition sites (whether type I, II or IIs) at which the sequences can be cut in a determinable fashion without loss of an essential biological function of the vector, and into which a nucleic acid fragment can be spliced in order to bring about its replication and cloning. Vectors can also comprise one or more recombination sites that permit exchange of nucleic acid sequences between two nucleic acid molecules. Such as, for example, subcloning of genes of interest between Entry and Destination vectors in the Gateway™ system (available from Invitrogen Corporation, Carlsbad, CA (see, e.g., Figure 36)). Vectors can further provide primer sites, e.g., for PCR, transcriptional and/or translational initiation and/or regulation sites, recombinational signals, replicons, Selectable markers, etc. Clearly, methods of inserting a desired nucleic acid fragment which do not require the use of recombination, transpositions or restriction enzymes (such as, but not limited to, UDG cloning of PCR fragments (U.S. Patent No. 5,334,575, entirely incorporated herein by reference), TA Cloning® brand PCR cloning (Invitrogen Corporation, Carlsbad, CA) (also known as direct ligation cloning), and the like) can also be applied to clone a fragment into a cloning vector to be used according to the present invention. The cloning vector can further contain one or more selectable markers suitable for use in the identification of cells transformed with the cloning vector.

[0117] Incorporating: As used herein, “incorporating” means becoming a part of a nucleic acid (e.g., DNA) molecule or primer.

- [0118] Nucleotide: As used herein, a "nucleotide" is a base-sugar-phosphate combination. Nucleotides are monomeric units of a nucleic acid molecule (DNA and RNA). The term nucleotide includes ribonucleoside triphosphates ATP, UTP, CTG, GTP and deoxyribonucleoside triphosphates such as dATP, dCTP, dITP, dUTP, dGTP, dTTP, or derivatives thereof. The term nucleotide as used herein also refers to dideoxyribonucleoside triphosphates (ddNTPs) and their derivatives. Illustrated examples of dideoxyribonucleoside triphosphates include, but are not limited to, ddATP, ddCTP, ddGTP, ddITP, and ddTTP. According to the present invention, a "nucleotide" may be unlabeled or detectably labeled by well known techniques. Detectable labels include, for example, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels and enzyme labels.
- [0119] Portion: As used herein, the term "portion" refers to part, or percentage of a whole entity. For example, a "portion" of a nucleic acid molecule refers to 1%, 10%, 25%, 50%, 75%, 90%, 99%, etc., of the whole nucleic acid molecule.
- [0120] Segment: As used herein, the term "segment" refers to part, or percentage of a whole entity. For example, a "segment" of a nucleic acid molecule refers to 1%, 10%, 25%, 50%, 75%, 90%, 99%, etc., of the whole nucleic acid molecule.
- [0121] Other terms used in the fields of recombinant nucleic acid technology and molecular and cell biology as used herein will be generally understood by one of ordinary skill in the applicable arts.
- [0122] The present invention relates to methods, compositions, isolated nucleic acids, vectors and kits for seamless cloning of nucleic acid molecules and production of nucleic acids and proteins.
- [0123] The vectors represented throughout, specifically shown in Figures 1A, 1B, 2A, 6A and 6B, 8A, 8B, 8C, 8D, 9A, 9B, 9C, 10, 11, 28, 33, 37 as well as similar vectors and portions of these vectors, may be used in the practice of the methods of the present invention. In each case, these vectors are designed such that upon digestion with a restriction enzyme (*e.g.* a type II restriction

enzyme), a sticky end is generated abutting and/or including nucleic acids which encode a peptide which may be cleaved from a protein or peptide encoded by a nucleic acid which is inserted into the vector. These, and other vectors of the present invention may further comprise one or more signal peptides and/or protease cleavage sites. The vectors of the present invention allow for the production of a protein that is exported from a cell and cleaved to generate a "mature" protein. The vectors of the present invention also allow for the production of a protein that is retained in the cell as a "native" protein.

[0124] In one aspect, the present invention provides methods for joining one or more (*e.g.* one, two, three, four, five, etc.) first nucleic acid molecules and a second one or more nucleic acid molecules, comprising: (a) combining the first and second nucleic acid molecules under conditions sufficient to allow for the joining of at least one terminus of the first nucleic acid molecule(s) to at least one terminus of the second nucleic acid molecule(s), wherein the terminus of the first nucleic acid molecule which is connected to the terminus of the second nucleic acid molecule(s) comprises a sticky end (*e.g.* an overhanging end) generated by a restriction enzyme (*e.g.* a type II restriction enzyme) and the terminus of the second nucleic acid molecule(s) is compatible (*e.g.* a blunt end or a sticky end) with this sticky end. In embodiments similar to the above and elsewhere herein, the sticky end may be on the terminus of the second nucleic acid molecule and the first nucleic acid molecule may contain a compatible end.

[0125] As in other embodiments of the invention described herein, the second nucleic acid molecule may contain an end which is generated by digestion with a type II restriction enzyme and the first nucleic acid molecule may contain a compatible end generated by other means.

[0126] In suitable embodiments, the present invention provides methods of cloning or subcloning one or more desired nucleic acid molecules comprising: (a) combining *in vitro* or *in vivo*, (i) one or more first nucleic acid molecules comprising one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* one or more type II restriction enzymes); and (ii)



one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more sticky ends on the first nucleic acid molecule(s) and, optionally, one or more selectable markers; and (b) incubating the combination under conditions sufficient to join the first nucleic acid molecule and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0127] In another aspect, the present invention provides methods for cloning or subcloning one or more desired nucleic acid molecules comprising: (a) combining *in vitro* or *in vivo*, (i) one or more first nucleic acid molecules comprising one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* one or more type II restriction enzymes); (ii) one or more second nucleic acid molecules comprising one or more restriction sites (*e.g.* one or more first type II restriction enzyme recognition sites) and, optionally, one or more selectable markers; and (iii) one or more restriction enzymes (*e.g.*, one or more type II restriction enzymes) that are specific for the restriction enzyme recognition site; and (b) incubating the combination under conditions sufficient to join the first nucleic acid molecule and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0128] In another aspect, the present invention provides methods for cloning or subcloning one or more desired nucleic acid molecules, or portions thereof, comprising: (a) combining *in vitro* or *in vivo*, (i) one or more first nucleic acid molecules comprising at least one nucleic acid segment that is flanked by one or more restriction sites (*e.g.* one or more first type II restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with a sticky end on the segment and, optionally, one or more selectable markers; and (iii) one or more restriction enzymes (*e.g.*, one or more type II restriction enzymes) that are specific for the restriction enzyme recognition site; and (b) incubating the combination under conditions sufficient to join the first nucleic acid segment and one or

more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0129] In another aspect, the present invention provides methods for cloning or subcloning one or more desired nucleic acid molecules, or portions thereof, comprising: (a) combining *in vitro* or *in vivo*, (i) one or more first nucleic acid molecules comprising at least one nucleic acid segment that is flanked by one or more first restriction sites (*e.g.* one or more first type IIs restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type IIs restriction enzyme recognition sites) and, optionally, one or more selectable markers; and (iii) one or more restriction enzymes (*e.g.* one or more type IIs restriction enzymes) that are specific for the first and/or second type IIs restriction enzyme recognition sites; and (b) incubating the combination under conditions sufficient to join the first nucleic acid segment and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0130] The seamless cloning methods of the present invention may utilize any restriction enzyme, including those which cleave nucleic acid molecules to produce blunt ends. The term "blunt ends" as used herein is used to indicate a nucleic acid molecule which has been cleaved by a restriction enzyme in such a way as to produce a double stranded nucleic acid in which both strands stop "bluntly" and do not overlap or overhang the other. Suitably, the methods of the invention utilize type IIs restriction sites. The present invention also encompasses the use of blunt-end cleavage enzymes, such as, but not limited to, *ScaI*, *SmaI*, *HpaI*, *HincII*, *HaeII* and *AluI*.

[0131] Type-IIs restriction enzymes and recognition sites which are useful in all aspects of the present invention include, but are not limited to, *EarI*, *MnII*, *PleI*, *AlwI*, *BbsI*, *BsaI*, *BsmAI*, *BspMI*, *Esp3I*, *HgaI*, *SapI*, *SfaNI*, *BbvI*, *BsmFI*, *FokI*, *BseRI*, *HphI*, *Alw26I*, *BbvII*, *BpmI*, *BsmI*, *BbsI*, *BsmBI*, *BaeI*, *BsrI*, *MlyI*, *BsrDI*, *Eco57I*, *GsuI*, *MnII*, *PleI*, *TaqII*, *Tth111II* and *MboII*. In all aspects of the present invention, the restriction enzyme recognition sites on the first and

second nucleic acid molecules may be the same sites or they may be different. In addition, the restriction enzyme recognition sites may be the same or different on each nucleic acid molecule. This allows for selective cloning where only nucleic acid segments with complementary sites will transfer between nucleic acids molecules.

[0132] Cleavage of a polynucleotide sequence with a type IIs restriction enzyme leaves an overhang on one strand of the sequence, or a sticky end. Via the cloning methods of the present invention, this sticky end can be combined with a compatible sequence on a second nucleic acid molecule resulting in a cloned, co-joined molecule. Sequences cleaved by Type IIs sites may also be joined to blunt ended compatible nucleic acid sequences via the cloning methods of the present invention. The compatible sequences can be joined via various catalyzing enzymes, for example DNA ligase and topoisomerase. Certain type IIs enzymes (*e.g.* *MlyI*) cleave and leave a blunt end on a nucleic acid molecule that may then be combined with a sticky end on a second nucleic acid molecule.

[0133] Nucleic acid molecules of the invention to be cloned may contain a blunt end to be linked, and the second nucleic acid molecule involved in the cloning method may contain an overhang at the end which is to be linked by a site-specific topoisomerase (*e.g.*, a type IA or a type IB topoisomerase), wherein the overhang includes a sequence complementary to that comprising the blunt end, thereby facilitating strand invasion as a means to properly position the ends for the linking reaction.

[0134] The nucleic acid molecules generated using this aspect of the invention include those in which one strand (not both strands) is covalently linked at the ends to be linked (*i.e.* double-stranded nucleic acid molecules generated using these methods contain a nick at each position where two ends were joined). These embodiments are particularly advantageous in that a polymerase can be used to replicate the double-stranded (ds) nucleic acid molecule by initially replicating the covalently linked strand. For example, a thermostable polymerase such as a polymerase useful for performing an amplification

reaction such as PCR can be used to replicate the covalently strand, whereas the strand containing the nick does not provide a suitable template for replication.

[0135] Preferably, the 5' termini of the ends of the nucleotide sequences to be linked by a type IB topoisomerase according to a method of certain aspects of the invention contain complementary 5' overhanging sequences, which can facilitate the initial association of the nucleotide sequences, including, if desired, in a predetermined directional orientation. Alternatively, the 5' termini of the ends of the nucleotide sequences to be linked by a type IB topoisomerase according to a method of certain aspects of the invention contain complementary 5' sequences wherein one of the sequences contains a 5' overhanging sequence and the other nucleotide sequence contains a complementary sequence at a blunt end of a 5' terminus, to facilitate the initial association of the nucleotide sequences through strand invasion, including, if desired, in a predetermined directional orientation. The term "5' overhang" or "5' overhanging sequence" is used herein to refer to a strand of a nucleic acid molecule that extends in a 5' direction beyond the terminus of the complementary strand of the nucleic acid molecule. Conveniently, a 5' overhang can be produced as a result of site specific cleavage of a nucleic acid molecule by a type IB topoisomerase.

[0136] Preferably, the 3' termini of the ends of the nucleotide sequences to be linked by a type IA topoisomerase according to a method of certain aspects of the invention contain complementary 3' overhanging sequences, which can facilitate the initial association of the nucleotide sequences, including, if desired, in a predetermined directional orientation. Alternatively, the 3' termini of the ends of the nucleotide sequences to be linked by a topoisomerase (e.g., a type IA or a type II topoisomerase) according to a method of certain aspects of the invention contain complementary 3' sequences wherein one of the sequences contains a 3' overhanging sequence and the other nucleotide sequence contains a complementary sequence at a blunt end of a 3' terminus, to facilitate the initial association of the nucleotide

sequences through strand invasion, including, if desired, in a predetermined directional orientation. The term "3' overhang" or "3' overhanging sequence" is used herein to refer to a strand of a nucleic acid molecule that extends in a 5' direction beyond the terminus of the complementary strand of the nucleic acid molecule. Conveniently, a 3' overhang can be produced upon cleavage by a type IA or type II topoisomerase.

[0137] The cloning methods of the present invention may be performed *in vitro* or *in vivo*. By *in vitro* and *in vivo* herein is meant cloning that is carried out outside of host cells (e.g., in cell-free systems, or in systems containing host cells in which the various cloning and recombination reaction(s) of the present invention take(s) place outside of the host cells) or inside of host cells (e.g., using recombination or other proteins expressed by host cells), respectively.

[0138] The nucleic acid molecules utilized and produced in the methods, compositions and kits of the present invention may be vectors or linear nucleic acid molecules. The term "vector," as used herein, refers to a nucleic acid molecule (preferably DNA) that provides a useful biological or biochemical property to an inserted nucleic acid. The terms "vector" and "plasmid" are used interchangeably herein. Examples of vectors include, phages, autonomously replicating sequences (ARS), centromeres, and other sequences which are able to replicate or be replicated *in vitro* or in a cell, or to convey a desired nucleic acid segment to a desired location within a cell of an animal. Vectors useful in the present invention include chromosomal-, episomal- and virus-derived vectors, e.g., vectors derived from bacterial plasmids or bacteriophages, and vectors derived from combinations thereof, such as cosmids and phagemids. A vector can have one or more restriction endonuclease recognition sites at which the sequences can be cut in a determinable fashion without loss of an essential biological function of the vector, and into which a nucleic acid fragment can be spliced in order to bring about its replication and cloning. Vectors can further provide primer sites, e.g., for PCR, transcriptional and/or translational initiation and/or regulation

sites, recombinational signals, replicons, selectable markers, etc. Clearly, methods of inserting a desired nucleic acid fragment which do not require the use of homologous recombination, transpositions or restriction enzymes (such as, but not limited to, UDG cloning of PCR fragments (U.S. Pat. No. 5,334,575, entirely incorporated herein by reference), TA Cloning® brand PCR cloning (Invitrogen Corp., Carlsbad, Calif.), and the like) can also be applied to clone a nucleic acid into a vector to be used according to the present invention. The vector can optionally further contain one or more selectable markers suitable for use in the identification of cells transformed with the vector, such as the selectable markers and reporter genes described herein. Vectors of the present invention may be derivative vectors as described throughout the present specification.

[0139] Vectors known in the art and those commercially available (and variants or derivatives thereof) may be used in the present invention. Such vectors may be obtained from, for example, Vector Laboratories Inc., Invitrogen, Promega, Novagen, NEB, Clontech, Boehringer Mannheim, Pharmacia, EpiCenter, OriGenes Technologies Inc., Stratagene, PerkinElmer, Pharmingen, and Research Genetics. General classes of vectors of particular interest include prokaryotic and/or eukaryotic cloning vectors, expression vectors, fusion vectors, two-hybrid or reverse two-hybrid vectors, shuttle vectors for use in different hosts, mutagenesis vectors, transcription vectors, vectors for receiving large inserts and the like.

[0140] Other vectors of interest include viral origin vectors (M13 vectors, bacterial phage  $\lambda$  vectors, adenovirus vectors, and retrovirus vectors), high, low and adjustable copy number vectors, vectors which have compatible replicons for use in combination in a single host (pACYC184 and pBR322) and eukaryotic episomal replication vectors (pCDM8).

[0141] Vectors for use in the present invention may comprise all, or portions of viral genomes, for example an adenovirus genome, a baculovirus genome, a herpesvirus genome, a pox virus genome, an adeno-associated virus genome,

a retrovirus genome, a flavivirus genome, a togavirus genome, an alphavirus genome, an RNA virus genome, etc.

[0142] The present invention also encompasses the use of recombinant retroviruses, e.g., lentiviruses, or any other type of retrovirus may be used in an analogous fashion to practice the present invention. A commercially available system for the construction of recombinant lentiviruses is ViraPower™ Lentiviral Expression System, available from Invitrogen Corporation, Carlsbad, CA. The ViraPower™ system provides a retroviral system for high-level expression in dividing and non-dividing eukaryotic cells, e.g., mammalian cells (See Figure 29). Examples of products available from Invitrogen Corporation, Carlsbad, CA include the ViraPower™ Lentiviral Directional TOPO® Expression Kit (catalog number K4950-00), the ViraPower™ Lentiviral GATEWAY™ Expression Kit (catalog number K4960-00), and the ViraPower™ Lentiviral Support Kit (catalog number K4970-00).

[0143] The present invention also encompasses replication-incompetent lentiviruses that can deliver and express one or more sequences of interest (e.g., genes). These viruses (based loosely on HIV-1) can effectively transduce dividing and non-dividing mammalian cells (in culture or *in vivo*), thus broadening the possible applications beyond those of traditional Moloney (MLV)-based retroviral systems (Clontech, Stratagene, etc.). Directional TOPO and GATEWAY™ lentiviral vectors have been created to clone one or more genes of interest with a V5 epitope, if desired. The Directional TOPO method involves a 5 minute bench-top ligation and results in 95% correct orientation (See Figures 33 and 34). The GATEWAY™ method involves cloning and sequencing a gene of interest only once into an entry clone and rapidly shuttling the gene of interest from vector to vector, or the destination clones. The GATEWAY™ method requires no restriction digests, gel purification or ligase. The GATEWAY™ method is 90-100% efficient and accurate and the gene of interest is cloned in the right direction and in-frame (Figure 35). The vectors also carry the blasticidin resistance gene (bsd) to

allow for the selection of transduced cells. Without additional modifications, these vectors can theoretically accommodate up to ~6 kb of foreign gene. Three supercoiled packaging plasmids (*gag/pol*, *rev* and VSV-G envelope) are provided to supply helper functions and viral proteins in trans (See Figures 30 and 32). Finally, an optimized producer cell line (293FT) is provided that will facilitate production of high titer virus. An Overview of lentiviral production is summarized in Figure 31 and involves the following steps: 1) Co-transfect 3 packaging plasmids and pLenti6-GOI into 293FT; 2) VSV-G envelope becomes studded in cell membrane; 3) *Rev* transports viral genome RNA with gene of interest out of the nucleus; 4) *gag* protein packages: viral RNA and *pol* protein; 5) Virus buds off cell, picks up envelope (pseudotyping). Plasmid maps of vectors adapted for use with GATEWAY™ and topoisomerase cloning in the production of nucleic acid molecules comprising all or a portion of a lentiviral genome are shown in Figures 8A (pLenti6/V5-DEST), 8B (pLenti6/V5-D-TOPO®), 8C (pLenti4/V5-DEST), and 8D (pLenti6/UbC/V5-DEST) respectively. The nucleotide sequences of the plasmids are provided in Tables 6-9, SEQ ID NOS:2-5. Plasmid maps of the three packaging plasmids pLP1, pLP2, and pLP/VSVG are shown in Figures 9A, 9B, and 9C respectively and the nucleotide sequences of these plasmids are provided as Tables 10, 11 and 12, (SEQ ID NOS:6-8) respectively.

[0144] Retroviruses are RNA viruses that reverse transcribe their genome and integrate the DNA copy into a chromosome of the target cell. It was discovered that the retroviral packaging proteins (*gag*, *pol* and *env*) could be supplied in trans, thus allowing the creation of replication incompetent viral particles capable of stably delivering a gene of interest. These retroviral vectors have been available for gene delivery for many years (Miller *et al.*, (1989) *BioTechniques* 7:980-990). One significant advantage of retroviral-based delivery is that the gene of interest is stably integrated into the genome of the host cell with very high efficiency. In addition, no viral genes are expressed in these recombinant vectors making them safe to use both in vitro and in vivo. However, one main drawback to the traditional Moloney-based



retroviruses is that the target cell must undergo one round of cell division for nuclear import and stable integration to occur. Traditional retroviruses do not have an active mechanism of nuclear import and therefore must wait for the host cell nuclear membrane to breakdown during mitosis before they can access the host genomic DNA (Miller *et al.*, *Mol. Cell. Biol.* 10:4239-442 (1990)).

[0145] Unlike traditional retroviruses, HIV (classified as a “lentivirus”) is actively imported into the nuclei of non-dividing cells (Lewis *et al.*, *J. Virol.* 68:510-516 (1994)). HIV still goes through the basic retrovirus lifecycle (RNA genome reverse transcribed in the target cell and integrated into the host genome); however, cis-acting elements facilitate active nuclear import, allowing HIV to stably infect non-dividing cells (for reviews see Buchschacher *et al.*, *Blood* 95:2499-2504 (2000), Naldini *et al.*, “The Development of Human Gene Therapy”, Cold Spring Harbor Laboratory Press, pages 47-60 (1999)). It is important to note that, for both lentivirus and traditional retroviruses, no gene expression occurs until after the viral RNA genome has been reverse transcribed and integrated into the host genome.

[0146] Similar to other retrovirus expression systems, the packaging functions of HIV can be supplied in trans, allowing the creation of lentiviral vectors for gene delivery. With all the viral proteins removed, the gene delivery vector becomes safe to use and allows foreign DNA to be efficiently packaged. In addition, it has been shown that lentiviral (or any retroviral) envelope proteins can be substituted for ones with broader tropism. The substitution of envelope is called pseudotyping, and allows creation of lentiviral vectors capable of infecting a wider variety of cells besides just CD4<sup>+</sup> cells. Many have found that the G protein from vesicular stomatitis virus (VSV-G) is an excellent pseudotyping envelope protein that imparts a very broad host range for the virus (Yee *et al.*, *Proc. Natl. Acad. Sci. USA* 91:9564-9568 (1994)). The ability of pseudo-typed lentivirus to infect a broad range of non-dividing cells has led to its extensive use in animal gene delivery and gene therapy (Baek *et*

*al.*, *Hum. Gene Ther.* 12:1551-8 (2001), Park et al., *Mol. Ther.* 4:164-73 (2001), Peng et al., *Gene Ther.* 8:1456-63 (2001)).

[0147] The present invention also encompasses the use of adenoviral vectors, including but not limited to, a pAd/PL-DEST vector (Table 11, Figure 10, SEQ ID NO:7) and pAd/CMV/V5-DEST vector (Table 12, Figure 11, SEQ ID NO:8). Adenoviruses are non-enveloped viruses with a 36 kb DNA genome that encodes more than 30 proteins. At the ends of the genome are inverted terminal repeats (ITRs) of approximately 100-150 base pairs. A sequence of approximately 300 base pairs located next to the 5'-ITR is required for packaging of the genome into the viral capsid. The genome as packaged in the virion has terminal proteins covalently attached to the ends of the linear genome.

[0148] The genes encoded by the adenoviral genome are divided into early and late genes depending upon the timing of their expression relative to the replication of the viral DNA. The early genes are expressed from four regions of the adenoviral genome termed E1-E4 and are transcribed prior to onset of DNA replication. Multiple genes are transcribed from each region. Portions of the adenoviral genome may be deleted without affecting the infectivity of the deleted virus. The genes transcribed from regions E1, E2, and E4 are essential for viral replication while those from the E3 region may be deleted without affecting replication. The genes from the essential regions can be supplied in trans to allow the propagation of a defective virus. For example, deletion of the E1 region of the adenoviral genome results in a virus that is replication defective. Viruses deleted in this region are grown on 293 cells that express the viral E1 genes from the genome of the cell.

[0149] In addition to permitting the construction of a safer, replication-defective viruses, deletion and complementation in trans of portions of the adenoviral genome and/or deletion of non-essential regions make space in the adenoviral genome for the insertion of heterologous DNA sequences. The packaging of viral DNA into a viral particle is size restricted with an upper limit of approximately 38 kb of DNA. In order to maximize the amount of

heterologous DNA that may be inserted and packaged, viruses have been constructed that lack all of the viral genome except the ITRs and packaging sequence (see, U.S. patent no. 6,228,646). All of the viral functions necessary for replication and packaging are provided in trans from a defective helper virus that is deleted in the packaging signal.

[0150] The present invention also encompasses the use of herpes viruses (see, for example, U.S. patent no. 5,672,344, issued to Kelly, *et al.*). The family Herpesviridae contains three subfamilies 1) alphaherpesvirinae, containing among others human herpesvirus 1; 2) betaherpesvirinae, containing the cytomegaloviruses; and 3) gammaherpesvirinae. Herpesviruses are enveloped DNA viruses. Herpesviruses form particles that are approximately spherical in shape and that contain one molecule of linear dsDNA and approximately 20 structural proteins. Numerous herpesviruses have been isolated from a wide variety of hosts. For example, United Patent No. 6,121,043 issued to Cochran, *et al.* describes recombinant herpesvirus of turkeys comprising a foreign DNA inserted into a non-essential region of the herpesvirus of turkeys genome; United States Patent No. 6,410,311 issued to Cochran, *et al.* describes recombinant feline herpesvirus comprising a foreign DNA inserted into a region corresponding to a 3.0 kb EcoRI-SalI fragment of a feline herpesvirus genome, United States Patent No. 6,379,967 issued to Meredith, *et al.*, describes herpesvirus saimiri, (HVS; a lymphotropic virus of squirrel monkeys) as a viral vector; and United States Patent No. 6,086,902 issued to Zamb, *et al.* describes recombinant bovine herpesvirus type 1 vaccines.

[0151] Herpesviruses have been used as vectors to deliver exogenous nucleic acid material to a host cell. In addition to the examples above, United States Patent No. 4,859,587, issued to Roizman describes recombinant herpes simplex viruses, vaccines and methods, United States Patent No. 5,998,208 issued to Fraefel, *et al.*, describes a helper virus-free herpesvirus vector packaging system, United States Patent No. 6,342,229 issued to O'Hare, *et al.*, describes herpesvirus particles comprising fusion protein and their preparation and use and United States Patent 6,319,703 issued to Speck describes

recombinant virus vectors that include a double mutant herpesvirus such as an herpes simplex virus-1 (HSV-1) mutant lacking the essential glycoprotein gH gene and having a mutation impairing the function of the gene product VP16.

[0152] Suitable vectors for use in the present invention also include prokaryotic vectors such as pcDNA II, pSL301, pSE280, pSE380, pSE420, pTrcHisA, B, and C, pRSET A, B, and C (Invitrogen, Corp.), pGEMEX-1, and pGEMEX-2 (Promega, Inc.), the pET vectors (Novagen, Inc.), pTrc99A, pKK223-3, the pGEX vectors, pEZZ18, pRIT2T, and pMC1871 (Pharmacia, Inc.), pKK233-2 and pKK388-1 (Clontech, Inc.), and pProEx-HT (Invitrogen, Corp.) and variants and derivatives thereof. Other vectors of interest include eukaryotic expression vectors such as pFastBac, pFastBacHT, pFastBacDUAL, pSFV, and pTet-Splice (Invitrogen), pEUK-C1, pPUR, pMAM, pMAMneo, pBI101, pBI121, pDR2, pCMVEBNA, and pYACneo (Clontech), pSVK3, pSVL, pMSG, pCH110, and pKK232-8 (Pharmacia, Inc.), p3'SS, pXT1, pSG5, pPbac, pMbac, pMC1neo, and pOG44 (Stratagene, Inc.), and pYES2, pAC360, pBlueBacHis A, B, and C, pVL1392, pBlueBacIII, pCDM8, pcDNA1, pZeoSV, pcDNA3 pREP4, pCEP4, and pEBVHis (Invitrogen, Corp.) and variants or derivatives thereof.

[0153] Other vectors suitable for use in the invention include pUC18, pUC19, pBlueScript, pSPORT, cosmids, phagemids, YAC's (yeast artificial chromosomes), BAC's (bacterial artificial chromosomes), P1 (*Escherichia coli* phage), pQE70, pQE60, pQE9 (quagan), pBS vectors, PhageScript vectors, BlueScript vectors, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene), pcDNA3 (Invitrogen), pGEX, pTrsfus, pTrc99A, pET-5, pET-9, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia), pSPORT1, pSPORT2, pCMVSPORT2.0 and pSV-SPORT1 (Invitrogen) and variants or derivatives thereof.

[0154] Additional vectors of interest include pTrxFus, pThioHis, pLEX, pTrcHis, pTrcHis2, pRSET, pBlueBacHis2, pcDNA3.1/His, pcDNA3.1(-)/Myc-His, pSecTag, pEBVHis, pPIC9K, pPIC3.5K, pAO815, pPICZ, pPICZ $\alpha$ , pGAPZ, pGAPZ $\alpha$ , pBlueBac4.5, pBlueBacHis2, pMelBac,

pSinRep5, pSinHis, pIND, pIND(SP1), pVgRXR, pcDNA2.1, pYES2, pZErO1.1, pZErO-2.1, pCR-Blunt, pSE280, pSE380, pSE420, pVL1392, pVL1393, pCDM8, pcDNA1.1, pcDNA1.1/Amp, pcDNA3.1, pcDNA3.1/Zeo, pSe, SV2, pRc/CMV2, pRc/RSV, pREP4, pREP7, pREP8, pREP9, pREP 10, pCEP4, pEBVHis, pCR3.1, pCR2.1, pCR3.1-Uni, and pCRBac from Invitrogen;  $\lambda$  ExCell,  $\lambda$  gt11, pTrc99A, pKK223-3, pGEX-1 $\lambda$ T, pGEX-2T, pGEX-2TK, pGEX-4T-1, pGEX-4T-2, pGEX-4T-3, pGEX-3X, pGEX-5X-1, pGEX-5X-2, pGEX-5X-3, pEZZ18, pRIT2T, pMC1871, pSVK3, pSVL, pMSG, pCH110, pKK232-8, pSL1180, pNEO, and pUC4K from Pharmacia; pSCREEN-1b(+), pT7Blue(R), pT7Blue-2, pCITE-4abc(+), pOCUS-2, pTA<sub>g</sub>, pET-32LIC, pET-30LIC, pBAC-2cp LIC, pBACgus-2cp LIC, pT7Blue-2 LIC, pT7Blue-2,  $\lambda$ SCREEN-1,  $\lambda$ BlueSTAR, pET-3abcd, pET-7abc, pET9abcd, pET11abcd, pET12abc, pET-14b, pET-15b, pET-16b, pET-17b-pET-17xb, pET-19b, pET-20b(+), pET-21abcd(+), pET-22b(+), pET-23abcd(+), pET-24abcd(+), pET-25b(+), pET-26b(+), pET-27b(+), pET-28abc(+), pET-29abc(+), pET-30abc(+), pET-31b(+), pET-32abc(+), pET-33b(+), pBAC-1, pBACgus-1, pBAC4x-1, pBACgus4x-1, pBAC-3cp, pBACgus-2cp, pBACsurf-1, plg, Signal plg, pYX, Selecta Vecta-Neo, Selecta Vecta-Hyg, and Selecta Vecta-Gpt from Novagen; pLexA, pB42AD, pGBT9, pAS2-1, pGAD424, pACT2, pGAD GL, pGAD GH, pGAD10, pGilda, pEZM3, pEGFP, pEGFP-1, pEGFP-N, pEGFP-C, pEBFP, pGFPUV, pGFP, p6xHis-GFP, pSEAP2-Basic, pSEAP2-Contral, pSEAP2-Promoter, pSEAP2-Enhancer, p $\beta$ gal-Basic, p $\beta$ gal-Control, p $\beta$ gal-Promoter, p $\beta$ gal-Enhancer, pCMV $\beta$ , pTet-Off, pTet-On, pTK-Hyg, pRetro-Off, pRetro-On, pIRES1neo, pIRES1hyg, pLXSN, pLNCX, pLAPSN, pMAMneo, pMAMneo-CAT, pMAMneo-LUC, pPUR, pSV2neo, pYEX4T-1/2/3, pYEX-S1, pBacPAK-His, pBacPAK8/9, pAcUW31, BacPAK6, pTriplEx,  $\lambda$ gt10,  $\lambda$ gt11, pWE15, and  $\lambda$ TriplEx from Clontech; Lambda ZAP II, pBK-CMV, pBK-RSV, pBluescript II KS +/-, pBluescript II SK +/-, pAD-GAL4, pBD-GAL4 Cam, pSurfscrip, Lambda FIX II, Lambda DASH, Lambda EMBL3, Lambda EMBL4,

SuperCos, pCR-Script Amp, pCR-Script Cam, pCR-Script Direct, pBS +/-, pBC KS +/-, pBC SK +/-, Phagescript, pCAL-n-EK, pCAL-n, pCAL-c, pCAL-kc, pET-3abcd, pET-11abcd, pSPUTK, pESP-1, pCMVLacI, pOPRSVI/MCS, pOPI3 CAT, pXT1, pSG5, pPbac, pMbac, pMC1neo, pMC1neo Poly A, pOG44, pOG45, pFRT $\beta$ GAL, pNEO $\beta$ GAL, pRS403, pRS404, pRS405, pRS406, pRS413, pRS414, pRS415, and pRS416 from Stratagene.

[0155] Two-hybrid and reverse two-hybrid vectors of interest include pPC86, pDBLeu, pDBTrp, pPC97, p2.5, pGAD1-3, pGAD10, pACt, pACT2, pGADGL, pGADGH, pAS2-1, pGAD424, pGBT8, pGBT9, pGAD-GAL4, pLexA, pBD-GAL4, pHISi, pHISi-1, placZi, pB42AD, pDG202, pJK202, pJG4-5, pNLexA, pYESTrp and variants or derivatives thereof.

[0156] The present invention also embodies the use and production of chimeric vectors. Such chimeric vectors may comprise one or more sequences that encode one or more functional or structural component of a viral vector, wherein each component may or may not come from the same or different types of viruses. Suitable components that may be combined to create such a chimeric vector include, but are not limited to, *gag*, *pol*, *env*, and *rev* genes and capsid proteins.

[0157] The nucleic acid molecules produced and/or utilized in the cloning methods, compositions and kits of the present invention may additionally or alternatively comprise one or more promoter molecules as described throughout the present specification, including the Pol III promoters H1 and U6 as well as other promoters recognized by RNA polymerase III. The nucleic acid molecules and vectors of the present invention may also further or alternatively comprise one or more genes which code for signal peptides and/or protease cleavage sites. Examples of protease cleavage sites include, but are not limited to, TEV sites and EK sites. TEV cleavage sites useful in the present invention include:

Consensus sequence: Glu-Xaa-Xaa-Try-Xaa-Gln//Xaa<sup>1</sup> (SEQ ID NO:23)

TEV1: Glu-Asn-Leu-Try-Phe-Gln//Xaa<sup>1</sup> (SEQ ID NO:24)

TEV2: Glu-Thr-Leu-Tyr-Ileu-Gln//Xaa<sup>1</sup> (SEQ ID NO:25)

(Xaa = any amino acid; Xaa<sup>1</sup> = any amino acid, except Pro; // = cleavage site).

[0158] EK cleavage sites useful in the present invention include:

Asp-Asp-Asp-Asp-Lys// (SEQ ID NO:26)

(// = cleavage site).

[0159] Signal peptides utilized in the present invention may be removed by a signal peptidase or any protease (*e.g.* Precision, thrombin and factor X) specific for one or more motifs on a signal peptide to generate a mature protein, including a protein encoded only by the inserted nucleic acid. The present invention also encompasses methods for the production of fusion proteins, and the fusion proteins produced by those methods. In accordance with the present invention, the proteins of the present invention may comprise one or more signal peptides, or portions of signal peptides, as noted above. These signal peptides may be used to facilitate production of desired proteins (*e.g.* mature or native proteins) *in vivo* or *in vitro*. Proteins produced using the methods of the present invention comprising such signal peptides would allow for the production of mature proteins, in which proteins are exported from the cell upon cleavage of the signal peptide by proteases within the cell. In an *in vitro* setting, these signal peptides would facilitate the production of native or desired proteins outside of a cell. Cleavage of the signal peptide may occur using signal peptidases, such as those described above, thus producing a desired protein product. These signal peptides may also be used as tags to facilitate affinity purification of polypeptides or proteins, for example fusion polypeptides or fusion proteins, produced by the methods of the present invention.

[0160] Any number of different protease recognition sites may be used in the practice of the invention. These sites will often be selected by to fit particular criteria suitable for the specific application. Exemplary proteases and protease recognition sites include the following. Tobacco Etch Virus (TEV) protease recognizes the amino acid sequence Glu-Xaa-Xaa-Tyr-Xaa-Gln//Xaa<sup>1</sup> (SEQ ID NO:23), where Xaa is any amino acid; Xaa<sup>1</sup> is any amino acid except Pro

and // indicates the cleavage site. Thus, for the amino acid sequence Glu-Asn-Leu-Tyr-Phe-Gln-Gly (SEQ ID NO:27), TEV cleaves between the Gln and Gly residues (see Invitrogen product literature associated with cat. nos. 10127-017 and 12575-015). Also, for the amino acid sequence Glu-Thr-Leu-Tyr-Ile-Gln-Xaa<sup>1</sup> (SEQ ID NO:25), TEV cleaves between the Gln and Xaa residues. Enterokinase (EK) recognizes the amino acid sequence Asp-Asp-Asp-Asp-Lys (SEQ ID NO:26) cleaves after the lysine (see Invitrogen product literature associated with cat. nos. E180-01 and E180-02, Invitrogen Corp., Carlsbad, CA). The ulp1 protease recognizes the amino acid sequence Gly-Gly-Ser (SEQ ID NO:28) and cleaves between the second glycine and the serine (U.S. Patent Publication No. 2003/0086918). Thus, the invention provides and includes nucleic acid molecules which may be used for producing proteins which may be processed by TEV protease, EK and/or ulp1 protease to generate proteins, as well as methods employing these enzymes and proteins or peptides produced using these methods.

[0161] In instances where the protein or peptide which is desired contains an amino terminal glycine, an amino terminal tag comprising and/or ending in a TEV protease recognition sequence may be used to generate a protein or peptides which contains no amino acids associated with, for example, cloning sites. Similarly, in instances where the protein which is desired contains an amino terminal serine, an amino terminal tag comprising and/or ending in a ulp protease recognition sequence may be used to generate a protein or peptide which contains amino acids associated with, for example, cloning sites. EK may be used to generate proteins or peptides which have an amino terminus other than glycine, as well as glycine.

[0162] The present invention also includes methods for joining two or more nucleic acid molecules using methods, for example, described elsewhere herein, wherein a first nucleic acid molecule contains a region which encodes a protease cleavage site and, optionally, a tag with a second nucleic acid molecule encodes a desired protein or peptide. In many instances, these nucleic acid segments are connected such that the desired protein is expressed



along with amino acids of the protease cleavage site as a fusion protein such that upon processing with the cognate protease, the desired protein is produced. Often, the desired protein which results from proteolytic digestion will contain only amino acids encoded by the second nucleic acid molecule referred to above.

[0163] In many instances, when a desired protein is produced from a nucleic acid formed by the connection of two nucleic acid molecules, the generation of a "seam" is only relevant with respect to one end of the protein (*i.e.*, the amino terminus or the carboxy terminus). In other words, in instances, where there is, for example, an amino terminal tag or a carboxy terminal tag, but not both, there is only a need to remove one tag. For example, when the translation product contains an amino terminal tag, the carboxy terminus of the translation product will typically terminate at a position in the mRNA which corresponds to the naturally resident stop codon. In such instances, a protease system may be used which will only remove amino terminal amino acids from the translation product.

[0164] The present invention also encompasses the production of a protein that comprises an expression enhancing amino acid sequence cleavable by ulp1 protease or an active fragment of ulp1 protease (for example the fragment from amino acid positions 403 to 621) and a poly-amino acid of interest, particularly one that is difficult to express in a recombinant expression system. The protein may also include a purification tag for ease of isolation. The ulp1 protease cleavable site may be any ulp1 cleavable site, such as for example a ulp1 protease cleavable site from a ubiquitin-like protein *e. g.* a SUMO (small ubiquitin-like molecule). The SUMO may be, for instance, Smt3 from yeast, or a fragment of Smt3 that retains the ability to be recognized and cleaved by Ulp 1. Examples of such a fragment of Smt3 include the fragment from amino acid positions 14-98 of Smt3 and the fragment from amino acid positions 1-98 of Smt3. Examples of such proteins can be found in WO 02/090495, the entire disclosure of which is incorporated herein by reference.

[0165] When nucleic acid molecules and/or methods of the invention are used to produce proteins or peptides, these proteins or peptides may be produced with an amino terminal and/or carboxy terminal tag. These tags may be used for any number of purposes, including to (1) increase the stability of the protein or peptide or (2) allow for purification. Thus, proteins or peptides produced by methods of the invention, as well as protein or peptides encoded by nucleic acid molecules of the invention, may contain affinity purification tags (*e.g.*, epitope tags such as the V5 epitope). Affinity purification tags are often amino acid sequences that can interact with a binding partner immobilized on a solid support. Nucleic acids encoding multiple consecutive single amino acids, such as histidine, may be used for one-step purification of the recombinant protein by affinity binding to a resin column, such as nickel sepharose. A protease cleavage site can be engineered between the affinity tag and the desired protein to allow for removal of the tag, for example, after the purification process is complete or to induce release of the desired protein or peptide from the solid support. Affinity tags which may be used in the practice of the invention include tags such as the chitin binding domain (which binds to chitin), polyarginine, glutathione-S-transferase (which binds to glutathione), maltose binding protein (which binds maltose), FLAsH, biotin (which binds to avidin and strepavidin), and the like.

[0166] Epitope tags are short amino acid sequences which are recognized by epitope specific antibodies. Proteins or peptides which contain one or more epitope tags may purified, for example, using a cognate antibody bound to a chromatography resin. The presence of the epitope tag furthermore allows the recombinant protein to be detected in subsequent assays, such as Western blots, without having to produce an antibody specific for the recombinant protein itself. Examples of commonly used epitope tags include V5, glutathione-S-transferase (GST), hemagglutinin (HA), the peptide Phe-His-His-Thr-Thr (SEQ ID NO:29), chitin binding domain, and the like. As discussed above, these affinity tags may be removed from the desired protein or peptide by proteolytic cleavage.

[0167] FLAsH tags comprise the sequence a cys-cys-Xaa-Xaa-cys-cys (SEQ ID NO:30), where Xaa and Xaa are amino acids. In many instances, Xaa and Xaa, which may be the same or different amino acids, are amino acids with high  $\alpha$ -helical propensity. In some embodiments, X and Y are the same amino acid. These peptides have been shown to bind to biarsenical compounds. The FLAsH systems is described in U.S. Patent No. 6,054,271, the entire disclosure of which is incorporated herein by reference.

[0168] The nucleic acid molecules and/or nucleic acid segments utilized in the cloning methods, compositions and kits of the present invention may optionally comprise one or more selectable markers comprising at least one DNA segment encoding an element selected from the group consisting of an antibiotic resistance gene, a gene that encodes a fluorescent protein, a tRNA gene, an auxotrophic marker, a toxic gene, a phenotypic marker, an antisense oligonucleotide, a restriction endonuclease, a restriction endonuclease cleavage site, an enzyme cleavage site, a protein binding site, and a sequence complementary to a PCR primer sequence.

[0169] Suitable antibiotic resistance genes for use in the present invention are well known in the art and include, but are not limited to, chloramphenicol resistance genes, ampicillin resistance genes, tetracycline resistance genes, Zeocin resistance genes, spectinomycin resistance genes and kanamycin resistance genes.

[0170] Examples of toxic gene products suitable for use in the present invention are well known in the art, and include, but are not limited to, restriction endonucleases (*e.g.*, DpnI), apoptosis-related genes (*e.g.* ASK1 or members of the bcl-2/ced-9 family), retroviral genes including those of the human immunodeficiency virus (HIV), defensins such as NP-1, inverted repeats or paired palindromic nucleic acid sequences, bacteriophage lytic genes such as those from ( $\Phi$ X174 or bacteriophage T4; antibiotic sensitivity genes such as rpsL, antimicrobial sensitivity genes such as pheS, plasmid killer genes, eukaryotic transcriptional vector genes that produce a gene product toxic to bacteria, such as GATA-1, and genes that kill hosts in the

absence of a suppressing function, *e.g.*, *kicB*, *sacB*, *ccdB*, ( $\Phi$ X174 E (Liu, Q. *et al.*, *Curr. Biol.* 8:1300-1309 (1998)), and other genes that negatively affect replicon stability and/or replication. The present invention also encompasses the use of a gene that encodes the *tus* gene which binds to one or more *ter* sites. A toxic gene can alternatively be selectable *in vitro*, *e.g.*, a restriction site.

[0171] Any of the nucleic acid molecules or nucleic acid segments used in or produced by the present methods, compositions and kits may further comprise one or more site-specific recombination sites. These recombination sites may flank the one or more restriction sites (*e.g.* one or more type II sites) if present in the nucleic acid molecules or segments of the invention. Site-specific recombinases are proteins that are present in or produced by many organisms (*e.g.*, viruses and bacteria) and have been characterized as having both endonuclease and ligase properties. These recombinases (along with associated proteins in some cases) recognize specific sequences of bases (*i.e.*, recombination sites) in a nucleic acid molecule and exchange the nucleic acid segments flanking those sequences. The recombinases and associated proteins are collectively referred to as "recombination proteins" (see, *e.g.*, Landy, A., *Current Opinion in Biotechnology* 3:699-707 (1993)).

[0172] Numerous recombination systems from various organisms have been described. See, *e.g.*, Hoess, *et al.*, *Nucleic Acids Research* 14:2287 (1986); Abremski, *et al.*, *J. Biol. Chem.* 261:391 (1986); Campbell, *J. Bacteriol.* 174:7495 (1992); Qian, *et al.*, *J. Biol. Chem.* 267:7794 (1992); Araki, *et al.*, *J. Mol. Biol.* 225:25 (1992); Maeser and Kahnmann, *Mol. Gen. Genet.* 230:170-176 (1991); Esposito, *et al.*, *Nucl. Acids Res.* 25:3605 (1997). Many of these belong to the integrase family of recombinases (Argos, *et al.*, *EMBO J.* 5:433-440 (1986); Voziyanov, *et al.*, *Nucl. Acids Res.* 27:930 (1999)). Perhaps the best studied of these are the Integrase/att system from bacteriophage ((Landy, A. *Current Opinions in Genetics and Devel.* 3:699-707 (1993)), the Cre/loxP system from bacteriophage P1 (Hoess and Abremski (1990) In *Nucleic Acids and Molecular Biology*, vol. 4. Eds.: Eckstein and Lilley, Berlin-Heidelberg:

Springer-Verlag; pp. 90-109), and the FLP/FRT system from the *Saccharomyces cerevisiae* 2  $\mu$  circle plasmid (Broach, et al., Cell 29:227-234 (1982)).

[0173] Recombination sites are sections or segments of nucleic acid on the participating nucleic acid molecules that are recognized and bound by the recombination proteins during the initial stages of integration or recombination. For example, the recombination site for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Curr. Opin. Biotech.* 5:521-527 (1994). Other examples of recognition sequences include the *attB*, *attP*, *attL*, and *attR* sequences which are recognized by the recombination protein Int. *attB* is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region, while *attP* is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). See Landy, *Curr. Opin. Biotech.* 3:699-707 (1993). Suitable recombination sites for use in the present invention include, but are not limited to, *attB* sites, *attP* sites, *attL* sites, *attR* sites, *lox* sites, *psi* sites, *tnpI* sites, *dif* sites, *cer* sites, *frt* sites, and mutants, variants and derivatives thereof.

[0174] The present cloning methods also embody the use of nucleic acid molecules that include a DNA segment having one or more terminal 3'-deoxyadenosine monophosphate (dAMP) residues, as described in US Patent No. 5,487,933, herein incorporated entirely by reference. These DNA segments are generated by thermophilic polymerases during PCR amplification. Double-stranded nucleic acids are formed with a single overhanging 3'-AMP residue. Mixture of these molecules with a population of linear double-stranded DNA molecules with a single overhanging deoxythymidylate (dTMP) residue at one or both of the 3' termini of the DNA molecule allow for ligation of the 3'-dAMP containing nucleic acid molecules

and the 3'-dTTP-containing DNA molecules to produce recombinant molecules. This approach is commonly known to those in the art as "TA Cloning," compositions and methods for which are available from Invitrogen Corporation (Carlsbad, CA).

[0175] The present invention also encompasses the use of cloning methods known to those skilled in the art as RecA cloning. The RecA cloning protein efficiently coats singly-stranded DNA. In the presence of ATP, this RecA coated single-stranded DNA can form triple-stranded nucleoprotein complexes with homologous double-stranded DNA. This RecA driven strand invasion and annealing can lead to high efficiency capture of DNA containing regions of homology with single-stranded DNA probes. This system can be used to increase the efficiency of recombination between a circular plasmid DNA molecule and a linear DNA "insert." Such suitable methods of RecA cloning can be found in U.S. Patent Nos. 5,948,653, 6,074,853 and 6,200,812, the disclosures of each of which are hereby incorporated entirely by reference.

[0176] The present invention also encompasses the use of a method of cloning DNA molecules in cells comprising the steps: a) providing a host cell capable of performing homologous recombination, b) contacting in said host cell a first DNA molecule which is capable of being replicated in said host cell with a second DNA molecule comprising at least two regions of sequence homology to regions on the first DNA molecule, under conditions which favour homologous recombination between said first and second DNA molecules and c) selecting a host cell in which homologous recombination between said first and second DNA molecules has occurred.

[0177] In this method of the present invention, the homologous recombination suitably occurs via the recET mechanism, i.e. the homologous recombination is mediated by the gene products of the recE and the recT genes which are preferably selected from the *E. coli* genes recE and recT or functionally related genes such as the phage  $\lambda$  red $\alpha$  and red $\beta$  genes. In contrast to RecA cloning, the recET cloning system requires significantly fewer bases of homology for efficient recombination into the target molecule. These proteins

facilitate the homologous incorporation of a double-stranded DNA fragment into a circular plasmid.

[0178] A host cell suitable for this embodiment of the present invention is a bacterial cell, *e.g.* a gram-negative bacterial cell. Suitably the host cell is an enterobacterial cell, such as *Salmonella*, *Klebsiella* or *Escherichia*. Most preferably the host cell is an *Escherichia coli* cell. It should be noted, however, that this method of the present invention is also suitable for eukaryotic cells, such as fungi, plant or animal cells. Such suitable methods of recET cloning can be found in Zhang, Y. *et al.*, *Nature* 20:123-128 (1998), Muryers, J.P.P., *et al.*, *Nucl. Acids Res.* 27:1555-1557 (1999), and U.S. Patent Nos. 6,509,156 and 6,355,412, the disclosures of each of which are hereby incorporated entirely by reference.

[0179] The first nucleic acid molecule and/or segment, as well as the second nucleic acid molecule involved in the methods, compositions and kits of the present invention may further or alternatively comprise one or more topoisomerase recognition sites and/or one or more topoisomerases. In suitable embodiments, the topoisomerase recognition site(s), if present, may optionally be flanked by two or more recombination sites.

[0180] The term "flanked" as used herein is meant to indicate a spatial relationship wherein a restriction site (*e.g.* a type II<sub>S</sub> site) and/or recombination site are located to one side of a nucleic acid segment (gene, selectable marker, etc.). As described above, recombination sites may also flank restriction sites (*e.g.* type II<sub>S</sub> sites) utilized in the invention. In the situation where a nucleic acid segment is flanked by two or more recombination or recognition sites, each side of the nucleic acid segment may be flanked by one or more sites.

[0181] Topoisomerases are categorized as type I, including type IA and type IB topoisomerases, which cleave a single strand of a double stranded nucleic acid molecule, and type II topoisomerases (gyrases), which cleave both strands of a nucleic acid molecule. Type IA and IB topoisomerases cleave one strand of a nucleic acid molecule. Cleavage of a nucleic acid molecule by type IA

topoisomerases generates a 5' phosphate and a 3' hydroxyl at the cleavage site, with the type IA topoisomerase covalently binding to the 5' terminus of a cleaved strand. In comparison, cleavage of a nucleic acid molecule by type IB topoisomerases generates a 3' phosphate and a 5' hydroxyl at the cleavage site, with the type IB topoisomerase covalently binding to the 3' terminus of a cleaved strand. The topoisomerase recognition sites of the present invention, if present, may be recognized and bound by a type I topoisomerase, and suitably by a type IB topoisomerase. Type IB topoisomerases useful in the present invention include, but are not limited to eukaryotic nuclear type I topoisomerase and a poxvirus topoisomerase. The poxvirus topoisomerase useful in the present invention may be produced by or isolated from a virus including, but not limited to, vaccinia virus, Shope fibroma virus, ORF virus, fowlpox virus, molluscum contagiosum virus and *Amsacta morrei* entomopoxvirus (see Shuman, *Biochim. Biophys. Acta* 1400:321-337, 1998; Petersen et al., *Virology* 230:197-206, 1997; Shuman and Prescott, *Proc. Natl. Acad. Sci., USA* 84:7478-7482, 1987; Shuman, *J. Biol. Chem.* 269:32678-32684, 1994; U.S. Pat. No. 5,766,891; PCT/US95/16099; PCT/US98/12372,, each of which is incorporated herein by reference; see, also, Cheng *et al.*, *supra*, 1998). Suitable type IB topoisomerases include the nuclear type I topoisomerases present in all eukaryotic cells and those encoded by vaccinia and other cellular poxviruses (see Cheng *et al.*, *Cell* 92:841-850, 1998, which is incorporated herein by reference). The eukaryotic type IB topoisomerases are exemplified by those expressed in yeast, *Drosophila* and mammalian cells, including human cells (see Caron and Wang, *Adv. Pharmacol.* 29B,:271-297, 1994; Gupta *et al.*, *Biochim. Biophys. Acta* 1262:1-14, 1995, each of which is incorporated herein by reference; see, also, Berger, *supra*, 1998).

[0182] In suitable aspects of the present invention, the one or more optional selectable markers of the nucleic acids or segment used in or produced by the present invention may be flanked by one or more restriction sites (e.g. one or more type IIs sites) and/or one or more recombination sites.



[0183] In other suitable embodiments of the present invention, the first nucleic acid molecule or segment and/or the second nucleic acid molecule may not comprise a promoter. The present invention allows for transfer of a promoter element into a second nucleic acid molecule that may not comprise a promoter via seamless cloning. In this orientation, transcription of the second nucleic acid molecule from the promoter element located on the first nucleic acid molecule may proceed such that no additional sequences are transcribed between the promoter element and the start codon of the second nucleic acid molecule. The present invention also allows for seamlessly adding a first nucleic acid molecule or segment into a second nucleic molecule that contains a promoter element such that the first nucleic acid molecule or segment will subsequently be under the control of the promoter element.

[0184] Incubation conditions suitable for use in the methods of the present invention comprise incubation with sufficient amounts of DNA ligases and buffers. Such incubation conditions are described in Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982). The term sufficient amount as used herein means that the amount of DNA ligase(s) and buffer(s) present during the cloning and/or recombination reactions is such that these reactions proceed as designed. Suitable buffers include physiologic buffers such as, but not limited to, Tris-(hydroxymethyl)aminomethane-HCl TRIS®-HCl, Ethylenediaminetetraacetic acid (EDTA) disodium salt, saline, Phosphate Buffered Saline (PBS), N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid) (HEPES®), 3-(N-Morpholino)propanesulfonic acid (MOPS), 2-bis(2-Hydroxyethylene)amino-2-(hydroxymethyl)-1,3-propanediol (bis-TRIS®), potassium phosphate (KP), sodium phosphate (NaP), dibasic sodium phosphate ( $\text{Na}_2\text{HPO}_4$ ), monobasic sodium phosphate ( $\text{NaH}_2\text{PO}_4$ ), monobasic sodium potassium phosphate ( $\text{NaKHPO}_4$ ), magnesium phosphate ( $\text{Mg}_3(\text{PO}_4)_2 \cdot 4\text{H}_2\text{O}$ ), potassium acetate ( $\text{CH}_3\text{COOH}$ ), D(+)- $\alpha$ -sodium glycerophosphate ( $\text{HOCH}_2\text{CH}(\text{OH})\text{CH}_2\text{OPO}_3\text{Na}_2$ ) and other physiologic buffers known to those skilled in the art.

[0185] In additional embodiments of the present invention provides methods for cloning or subcloning one or more desired nucleic acid molecules comprising: (a) combining *in vitro* or *in vivo* (i) one or more first nucleic acid molecules comprising one or more sticky ends generated by one or more first restriction enzymes (*e.g.* one or more type II restriction enzymes); (ii) one or more second nucleic acid molecules comprising one or more toxic genes flanked by one or more second restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); and (iii) one or more restriction enzymes (*e.g.* one or more type II restriction enzymes) that are specific for the first and/or second restriction sites; and (b) incubating the combination under conditions sufficient to join the first nucleic acid molecule and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules. Cloning via such methods of the invention allows for selection of successfully cloned nucleic acid molecules where the toxic gene originally present in the second nucleic acid molecule has been removed and replaced with a desired nucleic acid sequence from the first nucleic acid molecule.

[0186] In other embodiments of the present invention provides methods for cloning or subcloning one or more desired nucleic acid molecules, or portions thereof, comprising: (a) combining *in vitro* or *in vivo* (i) one or more first nucleic acid molecules comprising at least one nucleic acid segment that is flanked by one or more first restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more toxic genes flanked by one or more second restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); and (iii) one or more restriction enzymes (*e.g.* one or more type II restriction enzymes) that are specific for the first and/or second restriction enzyme recognition sites; and (b) incubating the combination under conditions sufficient to join the first nucleic acid molecule and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules. As noted above, cloning via such methods of the invention

allows for selection of successfully cloned nucleic acid molecules where the toxic gene originally present in the second nucleic acid molecule has been removed and replaced with a desired nucleic acid sequence from the first nucleic acid molecule.

[0187] The present invention also provides methods for cloning or subcloning one or more desired nucleic acid molecules comprising: (a) combining *in vitro* or *in vivo* (i) one or more first nucleic acid molecules comprising one or more sticky ends that have been generated by one or more first restriction enzymes (*e.g.* one or more type IIs restriction enzymes); (ii) one or more second nucleic acid molecules comprising one or more toxic genes and one or more antibiotic resistance genes all flanked by one or more second restriction sites (*e.g.* one or more type IIs restriction enzyme recognition sites); and (iii) one or more restriction enzymes (*e.g.* one or more type IIs restriction enzymes) that are specific for the restriction enzyme recognition sites; and (b) incubating said combination under conditions sufficient to join the first nucleic acid molecule into and or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules. This embodiment allows for additional selective screening via selection, for example, of antibiotic resistant host cells.

[0188] The present invention also provides methods for cloning or subcloning one or more desired nucleic acid molecules, or portions thereof, comprising: (a) combining *in vitro* or *in vivo* (i) one or more first nucleic acid molecules comprising at least one nucleic acid segment flanked by one or more first restriction sites (*e.g.* one or more type IIs restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more toxic genes and one or more antibiotic resistance genes all flanked by one or more second restriction sites (*e.g.* one or more type IIs restriction enzyme recognition sites); and (iii) one or more restriction enzymes (*e.g.* one or more type IIs restriction enzymes) that are specific for the restriction enzyme recognition sites; and (b) incubating said combination under conditions sufficient to join the first nucleic acid molecule and one or more of the second

nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules. This embodiment allows for additional selective screening via selection, for example, of antibiotic resistant host cells.

[0189] Another embodiment of the invention provides a method for cloning or subcloning one or more desired nucleic acid molecules comprising: (a) combining *in vitro* or *in vivo* (i) one or more first nucleic acid molecules comprising one or more sticky ends that have been generated by one or more first restriction enzymes (*e.g.* one or more type II restriction enzymes); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type II restriction enzyme recognition sites) flanked by one or more recombination sites; and (iii) one or more restriction enzymes (*e.g.* one or more type II restriction enzymes) that are specific for the first and/or second restriction enzyme recognition sites; and (b) incubating said combination under conditions sufficient to join the first nucleic acid molecule and one or more of said second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules. Following cloning of the first nucleic acid molecule, the cloned portion of the sequence may be cloned into another nucleic acid molecule via, for example, recombination cloning as described below.

[0190] Another embodiment of the invention provides a method for cloning or subcloning one or more desired nucleic acid molecules, or portions thereof, comprising: (a) combining *in vitro* or *in vivo* (i) one or more first nucleic acid molecules comprising at least one nucleic acid segment flanked by one or more first restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type II restriction enzyme recognition sites) flanked by one or more recombination sites; and (iii) one or more restriction enzymes (*e.g.* one or more type II restriction enzymes) that are specific for the first and/or second restriction enzyme recognition sites; and (b) incubating said combination under conditions sufficient to join the first nucleic acid molecule and one or more of said

second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules. As noted above, following cloning of the first nucleic acid molecule, the cloned portion of the sequence may be cloned into another nucleic acid molecule via, for example, recombination cloning as described below.

[0191] The present invention also provides for a method for cloning or subcloning one or more desired nucleic acid molecules, or portions thereof, comprising: (a) combining *in vitro* or *in vivo* (i) one or more first nucleic acid molecules comprising at least one nucleic acid segment flanked by one or more first restriction sites (*e.g.* one or more type II restriction enzyme recognition sites) and further flanked by one or more recombination sites; (ii) one or more second nucleic acid molecules comprising one or more recombination sites; and (iii) one or more site-specific recombination proteins; and (b) incubating the combination under conditions sufficient to transfer the first nucleic acid molecule into one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules.

[0192] This method of the present invention allows for the transfer of a nucleic acid sequence flanked by one or more restriction sites (*e.g.* one or more type II sites) that is further flanked by one or more recombination sites into a second nucleic acid molecule via recombinational cloning. Recombinational cloning is described in detail in U.S. Patent Nos. 5,888,732 and 6,277,608 (incorporated herein entirely by reference in their entireties). Recombinational cloning as disclosed in U.S. Patent Nos. 5,888,732 and 6,277,608 describes methods for moving or exchanging nucleic acid segments using at least one recombination site and at least one recombination protein to provide chimeric DNA molecules. Suitable recombination proteins for use in the present invention include, but are not limited to Int, Cre, IHF, Xis, Fis, Hin, Gin, Cin, Tn3 resolvase, TndX, XerC and XerD.

[0193] The methods of the present invention may further comprise introducing the product nucleic acid into one or more host cells. Host cells

that may be used in any aspect of the present invention include, but are not limited to, bacterial cells, yeast cells, plant cells and animal cells. Preferred bacterial host cells include *Escherichia* spp. cells (particularly *E. coli* cells and most particularly *E. coli* strains DH10B, Stbl2, DH5, DB3 (deposit No. NRRL B-30098), DB3.1 (preferably *E. coli* LIBRARY EFFICIENCY7 DB3.1J Competent Cells; Invitrogen Corporation, Carlsbad, CA), DB4 and DB5 (deposit Nos. NRRL B-30106 and NNRL B-30107 respectively, see U.S. Application No. 09/518,188, filed March 2, 2000, the disclosure of which is incorporated by reference herein in its entirety), JDP682 and *ccdA*-over (See U.S. Provisional Application No 60/475,004, filed June 3, 2003, the disclosure of which is incorporated by reference herein in its entirety), *Bacillus* spp. cells (particularly *B. subtilis* and *B. megaterium* cells), *Streptomyces* spp. cells, *Erwinia* spp. cells, *Klebsiella* spp. cells, *Serratia* spp. cells (particularly *S. marcessans* cells), *Pseudomonas* spp. cells (particularly *P. aeruginosa* cells), and *Salmonella* spp. cells (particularly *S. typhimurium* and *S. typhi* cells). Preferred animal host cells include insect cells (most particularly *Drosophila melanogaster* cells, *Spodoptera frugiperda* Sf9 and Sf21 cells and *Trichoplusia* High-Five cells), nematode cells (particularly *C. elegans* cells), avian cells, amphibian cells (particularly *Xenopus laevis* cells), reptilian cells, and mammalian cells (most particularly NIH3T3, CHO, COS, VERO, BHK and human cells). Preferred yeast host cells include *Saccharomyces cerevisiae* cells and *Pichia pastoris* cells. These and other suitable host cells are available commercially, for example from Invitrogen Corporation (Carlsbad, California), American Type Culture Collection (Manassas, Virginia), and Agricultural Research Culture Collection (NRRL; Peoria, Illinois).

[0194] Additional host cells that are useful in the present invention include mutant host cells and host cell strains, as well as mutants and/or derivatives thereof, that are resistant to the effects of the expression of one or more toxic genes. Host cells of this type may, for example, comprise one or more mutations in one or more genes within their genomes or on extrachromosomal or extragenomic DNA molecules (such as plasmids, phagemids, cosmids,

etc.), including mutations in, for example, *recA*, *endA*, *mcrA*, *mcrB*, *mcrC*, *hsd*, *deoR*, *tonA*, and the like, in particular in *recA* or *endA* or in both *recA* and *endA*. The mutations to these host cells may render the host cells and host cell strains resistant to toxic genes including, but not limited to, *ccdB*, *kicB*, *sacB*, *DpnI*, an apoptosis-related gene, a retroviral gene, a defensin, a bacteriophage lytic gene, an antibiotic sensitivity gene, an antimicrobial sensitivity gene, a plasmid killer gene, and a eukaryotic transcriptional vector gene that produces a gene product toxic to bacteria, and most particularly *ccdB*. Production and use of these type of mutant host cell strains are described in commonly owned U.S. Appl. Nos. 60/122,392, filed March 2, 1999, 09/518,188, filed March 2, 2000 (now abandoned), 10/396,696, filed March 20, 2003, and 60/475,004, filed June 3, 2003, the disclosures of which are incorporated herein by reference in their entireties.

[0195] Methods for introducing the cloned product nucleic acid molecules and/or vectors of the invention into the host cells described herein, to produce host cells comprising one or more of the cloned nucleic acid molecules and/or vectors of the invention, will be familiar to those of ordinary skill in the art. For instance, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells using well known techniques of infection, transduction, electroporation, transfection, and transformation. The nucleic acid molecules and/or vectors of the invention may be introduced alone or in conjunction with other the nucleic acid molecules and/or vectors and/or proteins, peptides or RNAs. Alternatively, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells as a precipitate, such as a calcium phosphate precipitate, or in a complex with a lipid. Electroporation also may be used to introduce the nucleic acid molecules and/or vectors of the invention into a host. Likewise, such molecules may be introduced into chemically competent cells such as *E. coli*. If the vector is a virus, it may be packaged *in vitro* or introduced into a packaging cell and the packaged virus may be transduced into cells. Hence, a wide variety of techniques suitable for introducing the nucleic acid molecules and/or vectors

of the invention into cells in accordance with this aspect of the invention are well known and routine to those of skill in the art. Such techniques are reviewed at length, for example, in Sambrook, J., *et al.*, Molecular Cloning, a Laboratory Manual, 2nd Ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 16.30-16.55 (1989), Watson, J.D., *et al.*, Recombinant DNA, 2nd Ed., New York: W.H. Freeman and Co., pp. 213-234 (1992), and Winnacker, E.-L., From Genes to Clones, New York: VCH Publishers (1987), which are illustrative of the many laboratory manuals that detail these techniques and which are incorporated by reference herein in their entireties for their relevant disclosures.

[0196] The present invention also encompasses producing a subsequent nucleic acid and/or a protein by introduction of a cloned product nucleic acid molecule of the invention and expression in a host cell. Methods and conditions by which to produce such product nucleic acid molecules and product proteins are well known in the art. See for example, Sambrook, J., *et al.*, Molecular Cloning, a Laboratory Manual, 2nd Ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (1989).

[0197] The present invention also encompasses the nucleic acid molecules and proteins produced from a host cell of the invention. An improvement of the present invention is that nucleic acid molecules produced using methods of the present invention, in many instances, will not contain extraneous nucleotides that are not associated with the desired nucleic acid, for example nucleotides encoded by the restriction sites (*e.g.* type II's restriction enzyme recognition sites). In other words, the seamless cloning methods of the present invention allow for a product molecule that does not contain extraneous nucleotides from other sources, including the restriction sites. Similarly, the product protein molecules produced using the methods of the present invention are free of amino acids that are not associated with the desired native or mature product protein, for example the product protein molecules are free of amino acids encoded by the restriction sites (*e.g.* type II's restriction sites). The proteins produced by the methods of the invention may be of any size,



including for example, a short peptide from about 5 amino acids, about 10 amino acids, about 20 amino acids, about 30 amino acids, about 40 amino acids, about 50 amino acids. The present invention also encompasses the production of larger proteins, for example about 300 amino acids in length, or even a large protein of greater than about 600 amino acids in length.

[0198] In one embodiment of the present invention, the nucleic acid molecules produced from the host cells may be useful as interfering RNA molecules. In biological systems that are not amenable to gene targeting or homologous recombination, a process called RNA interference (RNAi) is one practical method of generating knockout (KO) phenotypes. Post transcriptional gene silencing (PTGS) in plants and quelling in *Neurospora* was described in the early 1990s. RNAi was originally described in the model organism *C. elegans* as double stranded RNA (dsRNA) that mediated sequence specific gene silencing (Fire et al., *Nature* 391:806-811 (1998)). RNAi has also been described in yeast, *Drosophila*, plants and trypanosomes. RNAi can be used for genetic analysis. For example, it can be used for genome wide RNAi screens. RNAi has been shown to be conserved in mammals. RNAi has been used in the identification of a short interfering RNA (siRNA) as an effector molecule and with microRNA (miRNA) regulation. Essentially, the process involves application of double stranded RNA (dsRNA) that represents a complementary sense and anti-sense strand of a portion of a target gene within the region that encodes mRNA. The presence of the interfering dsRNA causes a severe post-transcriptional down-regulation of the target gene. This versatile technique has been used as a tool in the study of eukaryotic biology (see Sharp, P.A., *Genes Dev.* 13:139-141 (1999)). RNAi is an evolutionarily conserved phenomenon and a multi-step process that involves generation of active small interfering RNA (siRNA) *in vivo* through the action of an RNase III endonuclease, DICER, which digests long double stranded RNA molecules (dsRNA) into shorter fragments (See Figure 13). The 21- to 23-nucleotide base pair small interfering RNAs (siRNAs), produced through the action of DICER, mediate degradation of the complementary homologous RNA. One

bottleneck to using RNAi as a tool has been mRNA target site selection. Yet another challenge has been delivery, either transient such as transfection of dsRNA (See Figures 16-18)(Kawasaki et. al, NAR, 31(3):981-987 (2003)) or stable expression using vectors or a virus (See Figures 15 and 19)(Dykxhoorn, Novina and Sharp, Nature Reviews, Vol.4, (June 2003)). RNAi has successfully been reported in stable cell lines and transgenic mice. GFP shRNA block GFP expression in transgenic mice, decrease GFP in blastocytes and lower GFP fluorescence overall in a three day pup with two copies of the shRNA (Tiscornia et. al, PNAS, 2003).

[0199] RNAi is also powerful in reverse genetics. RNAi can be used as a loss of function tool, similar to antisense and ribozymes, but more potent. Natural cellular machinery use double stranded RNA to regulate cellular processes (e.g., miRNA). Some advantages of RNAi are that it is broadly conserved in eukaryotic organisms, is post transcriptional (effective in diploids) and is tunable (can adjust level of RNAi at several levels).

[0200] Until recently, RNAi technology did not appear to be applicable to mammalian systems. In mammals, dsRNA activates dsRNA-activated protein kinase (PKR) resulting in an apoptotic cascade and cell death (Der et al, *Proc. Natl. Acad. Sci. USA* 94:3279-3283 (1997)). In addition, it has long been known that dsRNA activates the interferon cascade in mammalian cells, which can also lead to altered cell physiology (Colby et al, *Annu. Rev. Microbiol.* 25:333 (1971); Kleinschmidt et al., *Annu. Rev. Biochem.* 41:517 (1972); Lampson et al., *Proc. Natl. Acad. Sci. USA* 58L782 (1967); Lomniczi et al., *J. Gen. Virol.* 8:55 (1970); Younger et al., *J. Bacteriol.* 92:862 (1966)). However, dsRNA-mediated activation of the PKR and interferon cascades typically require dsRNA longer than about 30 base pairs. Since the primary products of DICER are 21-23 base pair fragments of dsRNA, one can circumvent the adverse or undesired mammalian responses to dsRNA and still elicit an interfering RNA effect via siRNA (Elbashir et al., *Nature* 411:494-498 (2001)).

[0201] Thus, another aspect of the present invention provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules encoding one or more interfering RNAs that have one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* type II restriction enzymes); and (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more sticky ends on the first nucleic acid molecule(s), and optionally comprising one or more selectable markers; and (d) incubating the combination under conditions sufficient to join one or more of the nucleic acid molecules encoding the interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; (e) inserting the one or more product nucleic acid molecules into a host cell; and (f) expressing the one or more interfering RNAs in the host cell.

[0202] The present invention also provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules encoding one or more interfering RNAs flanked by one or more first restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type II restriction enzyme recognition sites) and optionally comprising one or more selectable markers; and (iii) one or more site-specific restriction enzymes (*e.g.* one or more type II restriction enzymes); and (d) incubating the combination under conditions sufficient to

join one or more of the nucleic acid molecules encoding the interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; (e) inserting the one or more product nucleic acid molecules into a host cell; and (f) expressing the one or more interfering RNAs in the host cell.

[0203] In yet another embodiment, the present invention provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules encoding one or more interfering RNAs that have one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* type II restriction enzymes); and (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more sticky ends on the first nucleic acid molecule(s), and optionally comprising one or more selectable markers; and (d) incubating the combination under conditions sufficient to join one or more of the nucleic acid molecules encoding the interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; and (e) expressing one or more interfering RNAs *in vitro* or *in vivo*. In a first further embodiment, the one or more interfering RNAs may be produced *in vitro* or isolated from a cell and then introduced into a second cell.

[0204] Another aspect of the present invention provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules encoding one or more interfering RNAs flanked by one or more first restriction sites (*e.g.* one or more type II restriction enzyme

recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type II's restriction enzyme recognition sites) and optionally comprising one or more selectable markers; and (iii) one or more site-specific restriction enzymes (*e.g.* one or more type II's restriction enzymes); and (d) incubating the combination under conditions sufficient to join one or more of the nucleic acid molecules encoding the interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; and (e) expressing one or more interfering RNAs *in vitro* or *in vivo*. In a first further embodiment, the one or more interfering RNAs may be produced *in vitro* or isolated from a cell and then introduced into a second cell.

[0205] Another aspect of the present invention provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules comprising one or more interfering RNAs that have one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* type II's restriction enzymes); and (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more sticky ends on the first nucleic acid molecule(s), and optionally comprising one or more selectable markers; and (d) incubating the combination under conditions sufficient to join one or more interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; (e) inserting the one or more product nucleic acid molecules into a host cell; and (f) expressing the one or more interfering RNAs in the host cell.

[0206] The present invention also provides methods of producing an RNA molecule for use as an interfering RNA comprising: (a) optionally, identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which comprise one or more interfering RNAs, wherein the

interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules comprising one or more interfering RNAs flanked by one or more first restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type II restriction enzyme recognition sites) and optionally comprising one or more selectable markers; and (iii) one or more site-specific restriction enzymes (*e.g.* one or more type II restriction enzymes); and (d) incubating the combination under conditions sufficient to join one or more interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; (e) inserting the one or more product nucleic acid molecules into a host cell; and (f) expressing the one or more interfering RNAs in the host cell.

[0207] Suitable nucleic acid molecules that can function as interfering RNA (iRNA) and that can be produced using the methods of the present invention may be either single- or double- stranded RNA (ssRNA or dsRNA, respectively). Examples of iRNA produced via methods of the present invention include, but are not limited to, antisense oligonucleotides, ribozymes, small interfering RNAs, double stranded RNAs, inverted repeats, short hairpin RNAs, small temporally regulated RNAs and the like.

#### Antisense Oligonucleotides

[0208] In general, antisense oligonucleotides comprise one or more nucleotide sequences sufficient in identity, number and size to effect specific hybridization with a preselected nucleic. Antisense oligonucleotides produced in accordance with the present invention typically have sequences that are selected to be sufficiently complementary to the target nucleic sequences (suitably mRNA in a target cell or organism) so that the antisense oligonucleotide forms a stable hybrid with the mRNA and inhibits the translation of the mRNA sequence, preferably under physiological conditions. It is preferred but not necessary that the antisense oligonucleotide be 100%

complementary to a portion of the target gene sequence. However, the present invention also encompasses the production of antisense oligonucleotides with a different level of complementarity to the target gene sequence, e.g., antisense oligonucleotides that are at least about 50% complementary, at least about 55% complementary, at least about 60% complementary, at least about 65% complementary, at least about 70% complementary, at least about 75% complementary, at least about 80% complementary, at least about 85% complementary, at least about 90% complementary, at least about 91% complementary, at least about 92% complementary, at least about 93% complementary, at least about 94% complementary, at least about 95% complementary, at least about 96% complementary, at least about 97% complementary, at least about 98% complementary, or at least about 99% complementary, to the target gene sequence.

[0209] Antisense oligonucleotides that may be produced in accordance with the present invention are well known in the art and that will be familiar to the ordinarily skilled artisan. Representative teachings regarding the synthesis, design, selection and use of antisense oligonucleotides include without limitation U.S. Patent No. 5,789,573, U.S. Patent No. 6,197,584, and Ellington, "Current Protocols in Molecular Biology," 2nd Ed., Ausubel *et al.*, eds., Wiley Interscience, New York (1992), the disclosures of which are incorporated by reference herein in their entireties.

#### Ribozymes

[0210] In general, ribozymes are RNA molecules having enzymatic activities usually associated with cleavage, splicing or ligation of nucleic acid sequences to which the ribozyme binds. Typical substrates for ribozymes include RNA molecules, although ribozymes may also catalyze reactions in which DNA molecules serve as substrates. Two distinct regions can be identified in a ribozyme: the binding region which gives the ribozyme its specificity through hybridization to a specific nucleic acid sequence, and a catalytic region which gives the ribozyme the activity of cleavage, ligation or splicing. Ribozymes

which are active intracellularly work in cis, catalyzing only a single turnover, and are usually self-modified during the reaction. However, ribozymes can be engineered to act in trans, in a truly catalytic manner, with a turnover greater than one and without being self-modified. Owing to the catalytic nature of the ribozyme, a single ribozyme molecule cleaves many molecules of target nucleic acids and therefore therapeutic activity is achieved in relatively lower concentrations than those required in an antisense treatment (WO 96/23569).

[0211] Ribozymes that may be produced in accordance with the present invention are well known in the art and that will be familiar to the ordinarily skilled artisan. Representative teachings regarding the synthesis, design, selection and use of ribozymes include without limitation U.S. Patent No. 4,987,071, and U.S. Patent No. 5,877,021, the disclosures of all of which are incorporated herein by reference in their entireties.

#### Small Interfering RNAs (siRNA)

[0212] RNAi is mediated by double stranded RNA (dsRNA) molecules that have sequence-specific homology to their "target" nucleic acid sequences (Caplen, N.J., et al., Proc. Natl. Acad. Sci. USA 98:9742-9747 (2001)). Biochemical studies in *Drosophila* cell-free lysates indicate that, in certain embodiments of the present invention, the mediators of RNA-dependent gene silencing are 21-25 nucleotide "small interfering" RNA duplexes (siRNAs). Accordingly, siRNA molecules are suitably used in methods of the present invention. The siRNAs are derived from the processing of dsRNA by an RNase known as Dicer (Bernstein, E., et al., *Nature* 409:363-366 (2001)). It appears that siRNA duplex products are recruited into a multi-protein siRNA complex termed RISC (RNA Induced Silencing Complex). Without wishing to be bound by any particular theory, a RISC is then believed to be guided to a target nucleic acid (suitably mRNA), where the siRNA duplex interacts in a sequence-specific way to mediate cleavage in a catalytic fashion (Bernstein, E., et al., *Nature* 409:363-366 (2001); Boutla, A., et al., *Curr. Biol.* 11:1776-1780 (2001)).



[0213] Small interfering RNAs that may be produced in accordance with the present invention are well known in the art and that will be familiar to the ordinarily skilled artisan. Small interfering RNAs that may be produced via the methods of the present invention suitably comprise between about 1 to about 50 nucleotides (nt). For example, siRNAs may comprise about 5 to about 40 nt, about 5 to about 30 nt, about 10 to about 30 nt, or about 15 to about 30 nt. Longer siRNAs (greater than about 30 nucleotides in length) may be useful in some non-human animal systems, and may suitably be produced by the methods of the present invention. Most reports describe the use of U6 or H1 pol III promoters to drive production of siRNA (Lee *et al.*, *Nat. Biotechnol.* 20:500-505 (2002); Paddison *et al.*, *Genes Dev.* 16:948-958 (2002); Brummelkamp *et al.*, *Science* 296:550-553 (2002)). Pol III promoters have all the elements required for initiation of transcription upstream of a defined transcription start site and terminate transcription at 4 or more Ts (incorporating only 1 or 2 Us into the 3' end of the nascent RNA). These attributes allow the production of short RNA molecules with defined ends.

#### Inverted Repeats

[0214] Inverted repeats comprise single stranded nucleic acid molecules that contain two sequences complementary to each other, oriented such that one of the sequences is inverted relative to the other. This orientation allows the two complementary sequences to base pair with each other, thereby forming a hairpin structure. The two copies of the inverted repeat need not be contiguous. There may be "n" additional nucleotides between the hairpin forming sequences, wherein "n" is any number of nucleotides. For example, n can be about 1, about 5, about 10, about 50, or about 100 nucleotide, or more, and can be any number of nucleotides falling within these discrete values.

[0215] Inverted repeats suitable that may be produced in accordance with the present invention can be synthesized and used according to procedures that are well known in the art and that will be familiar to the ordinarily skilled artisan. The production and use of inverted repeats for RNA interference can be found

in, without limitation, Kirby, K., *et al.*, *Proc. Natl. Acad. Sci. USA* 99:16162-16167 (2002), Adelman, Z. N., *et al.*, *J. Virol.* 76:12925-12933 (2002), Yi, C. E., *et al.*, *J. Biol. Chem.* 278:934-939 (2003), Yang, S., *et al.*, *Mol. Cell Biol.* 21:7807-7816 (2001), Svoboda, P., *et al.*, *Biochem. Biophys. Res. Commun.* 287:1099-1104 (2001), and Martinek, S. and Young, M. W., *Genetics* 156:171-1725 (2000).

#### Short Hairpin RNA (shRNA)

- [0216] Paddison, P.J., *et al.*, *Genes & Dev.* 16:948-958 (2002) have used small RNA molecules folded into hairpins as a means to effect RNAi. Accordingly, such short hairpin RNA (shRNA) molecules that may be produced via the methods of the present invention. Functionally identical to the inverted repeats described herein, the length of the stem and loop of functional shRNAs distinguishes them from inverted repeats. Stem lengths can range from about 1 to about 30 nt, and loop size can range between 1 to about 25 nt without affecting silencing activity. While not wishing to be bound by any particular theory, it is believed that these shRNAs resemble the dsRNA products of the Dicer RNase and, in any event, have the same capacity for inhibiting expression of a specific gene.
- [0217] Transcription of shRNAs is initiated at a polymerase III (pol III) promoter (e.g. U6 and H1 promoters) and is believed to be terminated at position 2 of a 4-5-thymine transcription termination site. Upon expression, shRNAs are thought to fold into a stem-loop structure with 3' UU-overhangs. Subsequently, the ends of these shRNAs are processed, converting the shRNAs into ~21 nt siRNA-like molecules.
- [0218] Short hairpin RNAs that may be produced in accordance with the present invention are well known in the art and that will be familiar to the ordinarily skilled artisan. The production and use of inverted repeats for RNA interference can be found in, without limitation, Paddison, P.J., *et al.*, *Genes & Dev.* 16:948-958 (2002), Yu, J-Y., *et al.* *Proc. Natl. Acad. Sci. USA* 99:6047-6052 (2002), and Paul, C. P. *et al.* *Nature Biotechnol.* 20:505-508 (2002).

#### MicroRNAs (miRNAs)

[0219] The invention may further be used to produce microRNA molecules. MicroRNA molecules are molecules which are structurally similar to shRNA molecules but, typically, contain one or more mismatches or insertion/deletions in their regions of sequence complementary. Hundreds of miRNAs have been identified in *C. elegans*, flies and humans. *C. elegans* miRNA, *lin-4* and *let-7*, have been identified to regulate developmental timing and inhibit expression of targeted genes. Examples of miRNA regulation from yeast to humans includes regulation of chromatin structure in yeast and tumor suppressor genes in humans. At least some microRNA molecules are transcribed as polycistrons of about 400, which are then processed to RNA molecules of about 70 nucleotides. These double stranded 70 mers are then processed again, presumably by the enzyme Dicer, to two RNA molecules which are about 22 nucleotides in length and often have one or more (e.g., one, two, three, four, five, etc.) internal mismatches in their regions of sequence complementarity. (See Figure 25) (Lee et al., EMBO 21:4663-4670 (2002). The miRNA can enter a miRNA ribonucleoprotein particle (miRNP) similar to siRNA entering into the RISC protein complex (Figure 14) (Dykxhoorn, Novina and Sharp, Nature Reviews, Vol.4, (June 2003)). The binding of miRNA/siRNAs of perfect complementarity to a target results in mRNA degradation; single base mismatches can block translation. The invention also includes, for example, uses of microRNA molecules and nucleic acid molecules which encode microRNA molecules which are similar to the uses described herein for shRNA and non-hairpin double stranded RNA molecules.

#### Small Temporally Regulated RNAs (stRNAs)

[0220] Another group of small RNAs that may be produced via the methods of the present invention are the small temporally regulated RNAs (stRNAs). In general, stRNAs comprise from about 20 to about 30 nt (Banerjee and

Slack, *Bioessays* 24:119-129 (2002)), although stRNAs of any size are also suitable for use in accordance with the invention. Unlike siRNAs, stRNAs downregulate expression of a target mRNA after the initiation of translation without degrading the mRNA.

[0221] The nucleic acids used in accordance with the present invention can be conveniently and routinely made through the well-known technique of solid-phase synthesis. Equipment for such synthesis is sold by several vendors including, for example, Applied Biosystems (Foster City, Calif.). Other methods for such synthesis that are known in the art may additionally or alternatively be employed. It is well-known to use similar techniques to prepare oligonucleotides such as the phosphorothioates and alkylated derivatives. By way of non-limiting example, see, e.g., U.S. Patent No. 4,517,338, and 4,458,066; Lyer RP, *et al.*, *Curr. Opin. Mol. Ther.* 1:344-358 (1999); and Verma S, and Eckstein F., *Annual Rev. Biochem.* 67:99-134 (1998), the disclosures of all of which are incorporated herein by reference in their entireties.

[0222] The present invention also provides methods for the production of gene knockout/knockdown cells and cells lines, as well as genetically modified transgenic animals.

[0223] In such suitable embodiments, the present invention provides methods of regulating the expression of one or more genes in a cell or an animal using interfering RNA, comprising: (a) identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules encoding one or more interfering RNAs that have one or more sticky ends that have been generated by one or more restriction enzymes (e.g. type II restriction enzymes); and (ii) one or more second nucleic acid molecules comprising one or more ends which are compatible with the one or more sticky ends on the first nucleic acid molecule(s), and optionally comprising one or more selectable markers; (d)

incubating the combination under conditions sufficient to join one or more of the nucleic acid molecules encoding the interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; and (e) inserting the one or more interfering RNA expression vectors into the cell or one or more cells of the animal, under conditions such that the one or more interfering RNAs bind to the one or more target nucleic acid sequences, thereby regulating expression of the one or more targeted genes.

[0224] The related embodiments, the present invention also provides methods of regulating the expression of one or more genes in a cell or an animal using interfering RNA, comprising: (a) identifying one or more target nucleic acid sequences; (b) preparing one or more nucleic acid molecules which comprise one or more interfering RNAs, wherein the interfering RNAs bind to the one or more target nucleic acid sequences; (c) combining *in vitro* or *in vivo*, (i) the one or more first nucleic acid molecules comprising one or more interfering RNAs flanked by one or more first restriction sites (*e.g.* one or more type IIs restriction enzyme recognition sites); (ii) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type IIs restriction enzyme recognition sites) and optionally comprising one or more selectable markers; and (iii) one or more site-specific restriction enzymes (*e.g.* one or more type IIs restriction enzymes); (d) incubating the combination under conditions sufficient to join one or more interfering RNAs and one or more of the second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules; and (e) inserting the one or more interfering RNA expression vectors into the cell or one or more cells of the animal, under conditions such that the one or more interfering RNAs bind to the one or more target nucleic acid sequences, thereby regulating expression of the one or more targeted genes.

[0225] The nucleic acid molecules of the invention can also be used to produce transgenic organisms (*e.g.*, animals and plants). Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs,

pigs, micro-pigs, goats, sheep, cows and non-human primates (e.g., baboons, monkeys, and chimpanzees) may be used to generate transgenic animals. Further, plants of any species, including but not limited to *Lepidium sativum*, *Brassica juncea*, *Brassica oleracea*, *Brassica rapa*, *Acena sativa*, *Triticum aestivum*, *Helianthus annuus*, Colonial bentgrass, Kentucky bluegrass, perennial ryegrass, creeping bentgrass, Bermudagrass, Buffalograss, centipedegrass, switch grass, Japanese lawnglass, coastal panicgrass, spinach, sorghum, tobacco and corn, may be used to generate transgenic plants.

[0226] Any technique known in the art may be used to introduce nucleic acid molecules of the invention into organisms to produce the founder lines of transgenic organisms. Such techniques include, but are not limited to, pronuclear microinjection (*Paterson et al., Appl. Microbiol. Biotechnol.* 40:691-698 (1994); *Carver et al., Biotechnology (NY)* 11:1263-1270 (1993); *Wright et al., Biotechnology (NY)* 9:830-834 (1991); and *Hoppe et al., U.S. Pat. No. 4,873,191* (1989)); retrovirus mediated gene transfer into germ lines (*Van der Putten et al., Proc. Natl. Acad. Sci., USA* 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (*Thompson et al., Cell* 56:313-321 (1989)); electroporation of cells or embryos (*Lo, Mol. Cell. Biol.* 3:1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e.g., *Ulmer et al., Science* 259:1745 (1993)); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (*Lavitrano et al., Cell* 57:717-723 (1989)); etc. For a review of such techniques, see *Gordon, "Transgenic Animals," Intl. Rev. Cytol.* 115:171-229 (1989), which is incorporated by reference herein in its entirety. Further, the contents of each of the documents recited in this paragraph is herein incorporated by reference in its entirety. See also, U.S. Patent No. 5,464,764 (*Capecchi et al., Positive-Negative Selection Methods and Vectors*); U.S. Patent No. 5,631,153 (*Capecchi et al., Cells and Non-Human Organisms Containing Predetermined Genomic Modifications and Positive-Negative Selection Methods and Vectors for Making Same*); U.S. Patent No. 4,736,866

(Leder *et al.*, Transgenic Non-Human Animals); and U.S. Patent No. 4,873,191 (Wagner *et al.*, Genetic Transformation of Zygotes); each of which is hereby incorporated by reference in its entirety.

[0227] Any technique known in the art may be used to produce transgenic clones containing nucleic acid molecules of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell *et al.*, *Nature* 380:64-66 (1996); Wilmut *et al.*, *Nature* 385:810-813 (1997)), each of which is herein incorporated by reference in its entirety).

[0228] The present invention provides for transgenic organisms that carry nucleic acid molecules of the invention in all their cells, as well as organisms which carry these nucleic acid molecules, but not all their cells, *i.e.*, mosaic organisms or chimeric. The nucleic acid molecules of the invention may be integrated as a single copy or as multiple copies such as in concatamers, *e.g.*, head-to-head tandems or head-to-tail tandems. The nucleic acid molecules of the invention may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko *et al.* (Lasko *et al.*, *Proc. Natl. Acad. Sci. USA* 89:6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that nucleic acid molecules of the invention be integrated into the chromosomal site of the endogenous gene, this will normally be done by gene targeting. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. Nucleic acid molecules of the invention may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu *et al.* (Gu *et al.*, *Science* 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation

will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. The contents of each of the documents recited in this paragraph is herein incorporated by reference in its entirety.

[0229] Once transgenic organisms have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze organism tissues to verify that integration of nucleic acid molecules of the invention has taken place. The level of mRNA expression of nucleic acid molecules of the invention in the tissues of the transgenic organisms may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the organism, in situ hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of tissue may which express nucleic acid molecules of the invention also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the expression product of these nucleic acid molecules.

[0230] Once the founder organisms are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular organism. Examples of such breeding strategies include, but are not limited to: outbreeding of founder organisms with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenic organisms that express nucleic acid molecules of the invention at higher levels because of the effects of additive expression of each copy of nucleic acid molecules of the invention; crossing of heterozygous transgenic organisms to produce organisms homozygous for a given integration site in order to both augment expression and eliminate the need for screening of organisms by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the nucleic acid molecules of the invention on a distinct background that is appropriate for an experimental model of interest.

[0231] Transgenic and "knock-out" organisms of the invention have uses which include, but are not limited to, model systems (e.g., animal model



systems) useful in elaborating the biological function of expression products of nucleic acid molecules of the invention, studying conditions and/or disorders associated with aberrant expression of expression products of nucleic acid molecules of the invention, and in screening for compounds effective in ameliorating such conditions and/or disorders.

[0232] As one skilled in the art would recognize, in many instances when nucleic acid molecules of the invention are introduced into metazoan organisms, it will be desirable to operably link sequences which encode expression products to tissue-specific transcriptional regulatory sequences (*e.g.*, tissue-specific promoters) where production of the expression product is desired. Such promoters can be used to facilitate production of these expression products in desired tissues. A considerable number of tissue-specific promoters are known in the art. Further, methods for identifying tissue-specific transcriptional regulatory sequences are described elsewhere herein.

[0233] The present invention also provides isolated nucleic acids comprising: (a) one or more sticky ends that have been generated by one or more restriction enzymes (*e.g.* one or more type II restriction enzymes); and (b) optionally one or more selectable markers. The present invention further provides isolated nucleic acids comprising: (a) one or more restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); and (b) optionally one or more selectable markers. As noted above, selectable markers for use in the isolated nucleic acids of the present invention comprise antibiotic resistance genes and toxic genes. As also described above, the isolated nucleic acids molecules of the present invention may also comprise one or more recombination sites, and one or more topoisomerase recognition sites and/or one or more topoisomerases. In suitable embodiments, the topoisomerase recognition site, if present, may optionally be flanked by two or more recombination sites.

[0234] In another embodiment, the present invention provides isolated nucleic acids comprising: (a) one or more sticky ends that have been generated by one

or more restriction enzymes (*e.g.* one or more type II restriction enzymes); and (b) one or more recombination sites. In yet another embodiment, the present invention provides isolated nucleic acids comprising: (a) one or more restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); and (b) one or more recombination sites. Suitable recombination sites include, but are not limited to, *attB* sites, *attP* sites, *attL* sites, *attR* sites, *lox* sites, *psi* sites, *mpI* sites, *dif* sites, *cer* sites, *frt* sites, and mutants, variants and derivatives thereof. In suitable embodiments, the isolated nucleic acid molecules of the present invention may optionally comprise one or more selectable markers, one or more topoisomerase recognition sites and/or one or more topoisomerases. In suitable embodiments, the topoisomerase recognition site, if present, may be flanked by two or more recombination sites. In additional embodiments, the one or more recombination sites may flank one or more restriction sites (*e.g.* one or more type II sites) and/or the one or more selectable markers, if present.

[0235] The present invention also provides vectors comprising: (a) one or more desired nucleic acid segments; (b) optionally one or more toxic genes; and (c) one or more restriction sites (*e.g.* one or more type II restriction enzyme recognition sites). Desired nucleic acid segments include, but are not limited to one or more genes, and one or more promoters. Suitable restriction sites include type II restriction enzyme recognition sites, such as those sites described above. The vectors of the present invention may also comprise one or more recombination proteins, and one or more topoisomerase recognition sites and/or one or more topoisomerases. In suitable embodiments, the topoisomerase recognition site, if present, may be flanked by two or more recombination sites. The vectors of the present invention optionally comprise suitable toxic genes, as described above. The vectors of the present invention may also optionally include one or more selectable marker as described throughout the specification. In another suitable embodiment, the vectors of the present invention may be "precut" by a restriction enzyme (*e.g.* a type II restriction enzyme). This precut vector may then be used to clone one more

second nucleic acid molecules which may comprise sticky ends, compatible with the vector, or optionally, may comprise on or more restriction sites (e.g. one or more type II's restriction enzyme recognition sites).

[0236] The present invention also provides methods of expressing and isolating nucleic acid molecules and proteins comprising: (a) obtaining one or more isolated nucleic acid molecules of the present invention; (b) introducing the isolated nucleic acid molecule into a host cell; (c) incubating the host cell under conditions sufficient to allow expression of a nucleic acid molecule or a protein encoded by the isolated nucleic acid molecule; and (d) isolating the expressed nucleic acid molecule or expressed protein. Host cells suitable for use in accordance with this aspect of the invention are described elsewhere herein. Suitable incubation conditions are well known in the art and are described in Freshney, R. I., "Culture of Animal Cells: A Manual of Basic Technique," Alan R. Liss, Inc, New York (1983) and Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982) and comprise incubating a host cell in a suitable growth medium with sufficient nutrients (e.g. Eagle's Minimum Essential Medium, DMEM: F12 Medium, RPMI-1640 Medium, Dulbecco's Modified Eagle's Medium, and the like) at an appropriate temperature (about 37°C). Methods of isolation of nucleic acid molecules and expressed proteins from host cells are also well known in the art and described in Maniatis *id.* and similar texts.

[0237] The expressed nucleic acid molecules may be suitable for use as interfering RNA as described above. As described throughout the specification, the expressed nucleic acid molecules will often not comprise extraneous, undesired nucleic acids, for example nucleic acids encoded by the one or more restriction sites (e.g. one or more type II's recognition sites). Similarly, the proteins produced via the methods of the present invention may not comprise extraneous, undesired amino acids, for example amino acids encoded by the one or more restriction sites (e.g. one or more type II's recognition sites).

[0238] The present invention also provides for methods of expressing desired nucleic acid segments comprising: obtaining a product nucleic acid molecule of the invention and incubating the nucleic acid molecule under conditions (*in vitro* or *in vivo*) such that the desired product nucleic acid molecule is transcribed and then translated. Incubation conditions for these methods of the invention are well known in the art as noted above.

[0239] The present invention also provides for methods of expressing desired nucleic acid segments comprising: (a) obtaining a vector of the present invention; (b) introducing the vector into a host cell; and (c) incubating the host cell under conditions sufficient to allow expression of a desired nucleic acid segment encoded by the vector. Incubation conditions for these methods of the invention are well known in the art as noted above.

[0240] Another embodiment of the present invention provides compositions comprising the elements described above that are involved in the various cloning methods of the invention. Such compositions comprise: (a) one or more first nucleic acid molecules comprising one or more sticky ends that have been generated by a restriction enzyme (*e.g.* one or more type II restriction enzymes); (b) one or more second nucleic acid molecules comprising one or more sticky ends which are compatible with the one or more sticky ends of the first nucleic acid molecule and, optionally, one or more selectable markers. Suitable restriction enzymes include those described throughout the specification, including, type II restriction enzyme recognition sites. The nucleic acids comprised in any of the compositions of the present invention may optionally further comprise one or more selectable markers, one or more recombination sites, one or more topoisomerase recognition sites and/or one or more topoisomerases and described above. The compositions may comprise one or more recombination proteins. Suitable recombination proteins include, but are not limited to, those described throughout the specification.

[0241] Another embodiment of the present invention provides compositions comprising the elements described above that are involved in the various

cloning methods of the invention. Such compositions comprise: (a) one or more first nucleic acid molecules comprising at least one nucleic acid segment flanked by one or more first restriction sites (*e.g.* one or more type II restriction enzyme recognition sites); (b) one or more second nucleic acid molecules comprising one or more second restriction sites (*e.g.* one or more type II restriction enzyme recognition sites) and optionally one or more selectable markers; and (c) one or more restriction enzymes (*e.g.* one or more type II restriction enzymes) that are specific for said first and/or second restriction enzyme recognition sites. Suitable restriction enzymes include those described throughout the specification, including, type II restriction enzyme recognition sites. The nucleic acids comprised in any of the compositions of the present invention may optionally further comprise one or more selectable markers, one or more recombination sites, one or more topoisomerase recognition sites and/or one or more topoisomerases and described above. The compositions may comprise one or more recombination proteins. Suitable recombination proteins include, but are not limited to, those described throughout the specification.

[0242] The present invention also provides kits comprising the isolated nucleic acids and/or vectors of the present invention. These kits are useful for practicing the various methods of the invention. Kits may comprise one or more first nucleic acid molecules and one or more second nucleic acid molecules. The first nucleic acid molecule may be an isolated nucleic acid molecule of the invention and the second nucleic acid molecule may be a vector of the present invention.

[0243] Kits of the invention may contain any number of components but typically will contain at least two components. Kits according to this aspect of the invention may comprise one or more containers, which may contain one or more components selected from the group consisting of one or more nucleic acid molecules or vectors of the invention, one or more primers, one or more polymerases, one or more reverse transcriptases, one or more recombination proteins, one or more restriction enzymes (*e.g.* one or more type II restriction

enzymes, or other enzymes for carrying out the methods of the invention), one or more topoisomerases, one or more buffers, one or more detergents, one or more restriction endonucleases, one or more nucleotides, one or more terminating agents (e.g., ddNTPs), one or more transfection reagents, pyrophosphatase, and the like. The kits of the invention may also comprise instructions for carrying out methods of the invention.

[0244] It will be readily apparent to one of ordinary skill in the relevant arts that other suitable modifications and adaptations to the methods and applications described herein may be made without departing from the scope of the invention or any embodiment thereof. Having now described the present invention in detail, the same will be more clearly understood by reference to the following examples, which are included herewith for purposes of illustration only and are not intended to be limiting of the invention.

## Examples

### Example 1

#### Expression of Interfering RNA using a Seamless Cloning Vector

[0245] The expression of short interfering hairpin RNA molecules (shRNA) *in vivo* can decrease the expression of genes with complementary sequences by RNA interference (RNAi) as described previously. The seamless cloning vector described here (pENTR/U6) allows for rapid and efficient cloning of double-stranded oligonucleotide pairs (~47bp) coding for a desired shRNA target sequence into a Pol III U6 expression cassette. The resulting shRNA vector contains an RNAi cassette flanked by *attL* sites. Therefore, the pENTR/U6 shRNA vectors can be used directly for transient transfection to test various shRNA target sequences, as well as to transfer the best shRNA cassettes to Lenti and Adenoviral DEST vectors for delivery into "hard to transfect" cells.

## Kit Components

[0246] Purified, *Bsa*I-linearized pENTR/U6.2 (once it is cut with *Bsa*I, i.e. the linear vector is called pENTR/U6) (Catalog No. K4945-00 and K4944-00, Invitrogen, Corp., Carlsbad, CA) Annealed lamin A/C control oligos: Top 5'-CACCGTGTTCTTCTGGAAGTCCAGCGAACTGGACTTCCAGAAGAACA (SEQ ID NO:9), Bottom 5'-AAAATGTTCTTCTGGAAGTCCAGTTCGCTGGACTTCCAGAAGAACA C (SEQ ID NO:10), Sequencing primers: U6 forward 5'-GGACTATCATATGCTTACCG (SEQ ID NO:11), M13 reverse 5'-CAGGAAACAGCTATGAC (SEQ ID NO:12)(Catalog No. N530-2, Invitrogen, Corp., Carlsbad, CA), T4 DNA ligase (Catalog No. 15224-025, Invitrogen, Corp., Carlsbad, CA) 5X T4 DNA ligase buffer (Catalog No. Y90001, Invitrogen, Corp., Carlsbad, CA), OneShot Top10 cells (Catalog No. C4040-03, Invitrogen, Corp., Carlsbad, CA). Thus, exemplary kits of the invention may comprise one, more, or all of these components.

## Vector Construction

[0247] *Entry vector.* The nucleic acid sequence of pENTR U6.2 (*Bsa*I-*ccdB*) is shown in Table 5, SEQ. ID. NO:1. The U6 promoter sequence was PCR amplified from genomic DNA (primers: 5'-AAGGTCGGG CAGGAAGAGGG-3' (SEQ ID NO:13); 5'-AGCGAGCACGGTGTTCGTC-3' (SEQ ID NO:14)) and TOPO cloned into pCR2.1/TOPO (included in kits, Catalog Nos. K4500-01, K4500-40, K4550-01, K4550-40, K4560-01, K4560-40, K4520-01 and K4520-40, Invitrogen, Corp., Carlsbad, CA). The promoter sequence was subsequently PCR amplified with the same primer sequences but with *Asp*718 and *Not*I sites appended to the primer 5' ends (5'GTGGGTACCAAGGTCGGGCAGGAAGAGGG-3' (SEQ ID NO:15; 5'-GTGGCGGCCGCGGTGTTTCGTCCTTTCCACAAG-3' (SEQ ID NO:16)). This PCR product was cloned by *Asp*718-*Not*I sticky end ligation into an

Entry vector with the pENTR/1a polylinker (Catalog No. 11813, Invitrogen, Corp., Carlsbad, CA) and pDONR/221 backbone (Catalog No. 12536-017, and provided in kits 12537-023, 12538-013, 12535-019, Invitrogen, Corp., Carlsbad, CA). The *ccdB* gene was amplified from pLenti6/V5/DEST (Catalog Nos. V496-10 and K4960-00, Invitrogen, Corp., Carlsbad, CA) (primers: 5'-GTGGCGGCCGCAAAGATCCTCCAGTGGATCCGGCTTACTAAAAG-3' (SEQ ID NO:17); 5'GTGCTCGAGAAAAAAGTCGACACGGAGCCCTCCAGTTATATTCCCCAGAACATCAGG-3' (SEQ ID NO:18)) and cloned into the above vector at the *NotI* and *XhoI* sites. These primers introduced *BpmI* restriction enzyme sites in the proper position at the ends of the PCR product and a 6bp polyT Pol III terminator.

[0248] To engineer the *BsaI* vector, a double stranded oligo containing a *BsaI* site and *NotI* site (5'GAGACCGCGGCCGCTTCTCGAGGTCTCATT (SEQ ID NO:19) + 5'TGAGACCTCGA GAAGCGGCCGCGGTCTCCG-3' (SEQ ID NO:20)) was cloned into *BpmI*-digested plasmid. The resulting plasmid was digested with *NotI* and *XbaI* and ligated to a new *ccdB* region PCR amplified (primers: 5'CACGCGGCCGCTGGATCCGGCTTACTAAAAG-3' (SEQ ID NO:21); 5'CACTCTAGAAAATGAGACCTTATATTCCCCAGAACATCAGG-3' (SEQ ID NO:22)) with a *NotI* site on one end and a *BsaI* site, 6bp polyT Pol III terminator, and *XbaI* site at the other. The final construct is named pENTR/U6.2 (*BsaI-ccdB*).

[0249] *LacZ* expression control vector. The *LacZ* expression control plasmid, pcDNA2.2' MS/GW/*LacZ* was made using Multi-site Gateway (CMVlacZV5). pENTR5'-CMV, pENTR-*LacZ* and pENTR/V5TKpolyA were mixed with the DEST R4R3 plasmid using LR Plus Clonase. The three plasmids in the Multi-site reaction were all created by a standard Gateway recombination reaction: 1) the CMV promoter was amplified from pcDNA3.1 (Catalog No. V790-20 and V795-20, Invitrogen, Corp., Carlsbad, CA) using primers flanked with *attB4* and *attB1* sequences and recombined with pDonr 5'(P4-P1R) to form pENTR5'-CMV. 2) The *LacZ* gene was amplified from



pcDNA3.1-*LacZ* using *attB1* and *attB2* flanking primers and recombined with pDonr 221 to create pENTR-*LacZ*, and, 3) the V5-TKpolyA element was amplified from pcDNA3.2 using *attB2* and *attB3* primers and recombined with pDonr3'(P2-P3R).

[0250]        *Preparation of linear pENTR/U6.2 ready for cloning.* pENTR/U6.2 in DB3.1 cells was grown in LB media with 50µg/ml kanamycin. Plasmid DNA was purified by SNAP midi prep with a yield of 67µg/50ml of culture. Ten µg of vector was digested with *BsaI* at 50°C in 200µl with 5 units of *BsaI*/µg of DNA for 2 hrs. After addition of 1.5vol of SNAP miniprep binding buffer, the reaction was added to a SNAP miniprep column, washed according to the SNAP protocol for miniprep DNA, and eluted in 100µl ddH<sub>2</sub>O and stored at – 20°C.

[0251]        *ShRNA Oligonucleotide Annealing.* DNA oligonucleotides of 46-53nt were produced with desalt purification only. Individual oligos were diluted in ddH<sub>2</sub>O to a final concentration of 200µM as verified by spectrophotometric analysis at OD<sub>260</sub>. Complementary oligos were mixed to the final desired concentration with either: 1) TE (10mM Tris pH 8.0, 1mM EDTA), 2) 10X Annealing Buffer and ddH<sub>2</sub>O such that the final, 1X buffer was 10mM Tris pH 8.0, 100mM NaCl, 1mM EDTA, or 3) the same buffer as in 2 but with a final concentration of 10mM MgCl<sub>2</sub>. (For example, to create a 50µM stock of a ds-oligo in 20µl, 5µl of each 200µM ss complementary oligo was mixed with 2µl of 10X Annealing buffer and 8µl of ddH<sub>2</sub>O). Mixed oligo pairs were heated and cooled in either an MJ thermocycler (94°C for 2 min, then decreased by 0.1°C every second to 25°C, and stored at 4°C) or incubation in a 95°C bath for 4min, then cooling to room temperature over 15min before putting the sample on ice. Annealed ds-oligos were diluted to the desired concentration with TE at room temperature.

[0252]        *Cloning target site DNA oligos into pENTR/U6.* *BsaI* cut pENTR/U6.2 and ds-oligos were incubated in a 20µl reaction using 5 times ligase buffer and 1µl ligase for 5 min at room temperature. Two microliters of the ligation

reaction were added to chemically competent Top10 One Shot cells (Catalog Nos. C4040-10, C4040-03, C4040-06, Invitrogen, Corp., Carlsbad, CA, ~50µl), incubated on ice for 20min, heat shocked at 42°C for 30 sec., and placed back on ice, followed by the addition of 250µl SOC and incubation at 37°C (shaking) for 1 hr. Ten to one hundred microliters of this transformation reaction were plated on LB Kan (50µg/ml) agarose plates.

[0253] The number of colonies per plate was determined after an overnight incubation at 37°C. A supercoiled pUC19 (2µl of a 10pg/µl stock) transformation control was performed with each set of cells transformed; in this case the transformation efficiency is reported as number of colony forming units per microgram.

[0254] *Sequence analysis of pENTR/U6 shRNA target clones.* Plasmid DNA was isolated from pENTR/U6 clones using the SNAP mini prep kit (Catalog No. K1900-01, Invitrogen, Corp., Carlsbad, CA) under standard conditions. Two different primers were used for sequence analysis:

1) U6 forward, 5'-GGACTATCATATGCTTACCG (forward primer, binds in U6 promoter 55bp from the 3' end of the U6 promoter)(SEQ ID NO:11)

2) M13 R, 5'-CAGGAAACAGCTATGC (reverse primer, binds "downstream" from the AttL2 site, 146bp from the pol III termination)(SEQ ID NO:12)

[0255] *Gateway LxR recombination.* 150ng of each pENTR/U6 shRNA clone and 150ng of pLenti6/PL-DEST or 300ng of pAD/PL-DEST (Figure 10) (Catalog No. V494-20, Invitrogen, Corp., Carlsbad, CA) were incubated in a 20µl reaction using the 5X buffer and 5X LR Clonase enzyme mix, and incubated at 25°C for 1hr. Two microliters of this LxR reaction were transformed into chemically competent cells as described above except that selection plates had 50ug/ml ampicillin instead of kanamycin.

### ShRNA transfections

[0256] All transfections were carried out in 24-well plates. For luciferase and  $\beta$ -galactosidase ( $\beta$ -gal) knockdown experiments, 600ng of pENTR/U6-shRNA vectors were cotransfected with 100ng each pcDNA5/FRT/luc and the pcDNA1.2/V5-GW/lacZ positive control plasmid into GripTite™ 293 cells (Catalog No. R795-07, Invitrogen, Corp., Carlsbad, CA) using Lipofectamine 2000™. Briefly, cells were plated the day before transfection in 0.5ml medium lacking antibiotics at  $2 \times 10^5$  cells per well. On the day of transfection, cells were typically 90-95% confluent. For each well, 2 $\mu$ l of Lipofectamine 2000™ were diluted with 48 $\mu$ l OptiMEM, incubated 5 min at room temperature, then mixed with DNAs diluted with OptiMEM to 50 $\mu$ l. Complexes were incubated an additional 20min at room temperature before addition to cells. Medium was changed 3hr after transfection to minimize toxicity.

### Luciferase and $\beta$ -gal assays.

[0257] After 48hr, GripTite™ 293 cells were lysed in 0.5ml luciferase lysis buffer (25mM Tris-HCl pH 8.0, 0.1mM EDTA pH 8.0, 10% glycerol, 0.1% Triton X-100) and subjected to a -80°C freeze-thaw. 50 $\mu$ l of each lysate was used in a luciferase luminescence assay (Promega) while another 10 $\mu$ l was used in a  $\beta$ -gal luminescence assay (Tropix) according to the manufacturers' instructions.

### Results

[0258] The vector pENTR/U6 is designed to express shRNA in mammalian cells for use in RNAi. (pENTR/U6.2 is the supercoiled vector containing the *ccdB* gene; once linearized with *Bsa*I, the vector will be referred to as pENTR/U6.) pENTR/U6 allows the cloning of shRNA target sequences between the human U6 pol III promoter and a 6 T termination signal in a Gateway Entry (ENTR) vector. In this case, the entire RNAi cassette (U6

promoter, cloning site, and termination signals) is between the *attL1* and *attL2* recombination sites. Therefore, U6 driven expression of an shRNA is possible directly from ENTR vector and does not require subsequent LxR transfer to a DEST vector.

#### Vector preparation

[0259] pENTR/U6.2 (*BsaI*-*ccdB*) is digested with the type IIS restriction enzyme *BsaI* in preparation for cloning ds-oligos (~47mers) containing shRNA target sequences. Type IIS restriction enzymes cut outside of their recognition sequence and can therefore be used to create sticky ends of any sequence in the vector. In this case, the *BsaI* digest leaves the 4nt 5' ssDNA end 3'-GTGG-5' at the end of the U6 promoter and the single stranded 3'-TTTT-5' at the other vector end (the first four Ts of the termination signal).

[0260] Digestion of the pENTR/U6.2 by *BsaI* generates three fragments (2850, 577, and 91bp). The linearized cloning vector is 2850bp; smaller fragments derive from the *ccdB* gene (*ccdB* has a *BsaI* site). Removal of the smaller fragments from the final vector prep is not required; however, the amount of the 91bp fragment recovered from the SNAP purification can vary. Uncut pENTR/U6.2 or clones that have reassembled the functional *ccdB* gene will not propagate in Top10 cells. The cloning efficiency of either small fragment alone is very low due to non-compatible ends.

#### Insert Annealing

[0261] A five-minute bench top ligation and subsequent transformation is highly efficient at cloning dsDNA oligo shRNA target sequences - if the oligo inserts are properly annealed. A typical 46nt ss-oligo is made of a 4nt 5' cloning overhang followed by 19nt of "sense" and a complementary 19nt "antisense" sequence connected by short 4nt "loop." Thus the oligos can form a ~19bp DNA intra-molecular hairpin. Therefore, conditions must be optimized to favor intermolecular annealing between two different complementary oligos rather than the production of single-strand

intramolecular hairpins. The formation of intermolecular ds-oligos can be accomplished by melting (heating to 94°C) and cooling complementary oligos at high concentrations in the appropriate buffer.

[0262] Intermolecular double-stranded molecules can be formed in annealing buffers containing either 20 or 100mM NaCl when the oligo concentration is 50µM during the heating and cooling cycle. The ds-molecules can be separated from the single-stranded hairpins in an E-gel. Additionally, no difference was noted between using the Thermocycler or water bath protocols to melt/cool the reaction.

[0263] Upon closer examination of the salt and oligo concentration, a buffer without any NaCl (TE) would not support formation of ds-47mers even at 100µM concentrations, adding MgCl<sub>2</sub> to 100mM NaCl had no effect, and oligo concentrations of less than 50µM were compromised in the amount of ds-47mers created.

[0264] Once created, the dsDNA 47mer shRNA inserts can be diluted in TE for cloning. After the ds-47mers are diluted, they are stable at 4°C overnight, but will form single strand hairpins if melted, i.e. incubated at temps above 42°C.

[0265] Heating and cooling of shRNA target oligos at concentrations of 50µM or greater in 10mM Tris pH 8.0, 100mM NaCl, 1mM EDTA creates a mixture of ~50:50 ds/hairpin molecules which can be effectively cloned into *Bsa*I linearized pENTR/U6 (see pENTR/U6 cloning, below).

#### Gateway ENTR vector testing

[0266] The supercoiled pENTR/U6.2 (*Bsa*I-*ccdB*) vector, prior to linearization for cloning, passes the criteria set for Gateway ENTR vectors (> 10<sup>4</sup> killing by *ccdB*). Supercoiled pENTR/U6.2 was transformed into *E. coli* cells it should kill (Top10 and HB101 cells) as well as the DB3.1 cell line designed to propagate plasmids with the *ccdB* gene. pENTR/U6.2 transforms DB3.1 cells 1.3 x 10<sup>4</sup> times better than Top10s cells once the number of

colonies per plate are adjusted for the different transformation efficiencies of the different cell lines (the Top10 cells were ~200 times more competent than the DB3.1 cells and ~400 times more competent than the HB101 cells).

[0267] When *Bsa*I digestion of pENTR/U6.2 is complete, most of the supercoiled vector is linearized. Transformation of *Bsa*I cut, SNAP purified pENTR/U6 vector only generated a small number of “background” colonies per plate in Top10 or DB3.1 cells. Eight colonies were obtained in DB3.1 cells and all looked like the parent sc pENTR/U6.2 by RFLP analysis (data not shown) indicating the *Bsa*I digest is efficient and only a small fraction of the plasmids are left uncut after the 2hr incubation. In Top10 cells only 4 colonies were obtained; RFLP analysis of these indicated two classes, neither of which was the parent plasmid (possibly pENTR/U6 closed without the *ccdB* gene and one fragment of the *ccdB* gene re-cloned).

#### pENTR-U6 cloning

[0268] A five-minute bench-top ligation is an easy and efficient method to clone shRNA target sequences into pENTR/U6. The cloning process was optimized over a wide range of vector concentrations (20pg – 5ng) and insert concentrations (0.4pg – 10ng) with the shRNA target sequence lacZ-19. All the optimization of the cloning reaction was done with ds-oligos annealed at a concentration of 50μM prior to dilution in TE and transformation into chemically competent Top10 cells. Sequence analysis of the shRNA clones demonstrate that >90% have inserts in the correct orientation.

[0269] Greater than 15 other ds-oligo inserts, each with a different shRNA target sequence, have been cloned into pENTR/U6 under comparable conditions. In all cases, the number of colonies generated was similar to the numbers of colonies generated with the lacZ-19 ds-oligo. No significant difference has been noted in how different inserts clone into the pENTR/U6 vector.

### Sequence analysis

[0270] The efficiency of cloning shRNA target-sequence inserts was determined by sequence analysis through shRNA target sequences. Analysis of the lacZ-19 shRNA target inserts cloned in pENTR/U6 under the recommended conditions, demonstrated that 100% (38/38) of the randomly selected clones have an insert cloned in the correct orientation.

[0271] Sequence analysis with the U6 forward primer provides excellent sequence through the cloned shRNA target sequence. It is designed for ease of analysis of the cloned oligos, binds the U6 promoter inside the *attL* sites 55 bases from the cloning junction, and allows for the analysis of the entire cloned insert with a 100 base “read” before the “downstream” *attL2* site.

### RNAi by transient Transfections

[0272] Post-transcriptional inhibition of luciferase (GL2) and lacZ expression was evident upon expression of shRNA targets from the pENTR/U6 vector (Figure 3A). Specific inhibition is evident with pENTR/U6 shRNA clones targeting Luciferase and lacZ expression from co-transfected reporter constructs. The Luciferase pENTR/U6 GL2-22 construct inhibits expression of GL2 Luciferase but not lacZ (Figure 3A); similarly, the pENTR/U6 with the lacZ-19 shRNA target sequence (the target provided as a control in this kit) inhibits lacZ expression from pcDNA1.2/V5-GW/lacZ (the control expression vector for this kit) - but not Luciferase (Figure 3B).

[0273] Similar inhibition of both lacZ and Luciferase is evident with shRNAs that target different sites, although not all shRNA sequences are effective (Figures 4A and 4B). The kit control lacZ-19 target site presented in Figure 4B is the same shRNA target site used in Figure 3B, and only the lacZ4-AS sequence inhibits expression to the same degree. The lacZ4-SA only moderately inhibits expression and the lacZ5 clones have little if any inhibitory effect. Similarly, the GL2sh2 and GL2-22 (AS) target sites are the most effective shRNA clones tested at inhibiting luciferase expression

(Figures 4A and 4B). Interestingly, the sense to anti-sense orientation of the shRNA target sequence can make a considerable difference in the level of inhibition at a specific target (Figures 4A and 4B). However, the optimal orientation (sense-loop-antisense (SA) or antisense-loop-sense (AS)) is not clear; with Luciferase, the AS orientation was most effective, but with lacZ the SA orientation was most effective (Figure 4A, ENTR/U6-A6-GL2-22 AS vs. SA, and Figure 4B, ENTR/U6-A6-lacZ4-AS vs. SA).

[0274] Additionally, the lacZ-19 shRNA target sequence was tested in derivatives of the pENTR/U6 vector with terminators of 4-8 Ts. All the terminators behaved similarly (Figure 5).

#### Gateway LxR cross

[0275] Any shRNA target sequence cloned into pENTR/U6 can easily be transferred as a U6 RNAi cassette to a Gateway DEST vector by attL x attR (LxR) recombination at the *att* sites. Following is a demonstration of the efficiency of LxR transfer. The lacZ-19 target sequence cloned into pENTR/U6 was transferred into pLenti6/PL-DEST and pAD/PL-DEST by a standard LxR Clonase catalyzed recombination reaction (See, e.g., Figs. 38 and 39) as described previously (See U.S. Patent Nos. 5,888,732; 6,143,577; 6,171,861; 6,277,608; and 6,720,140; the disclosures of which are incorporated by reference herein in their entireties). Additionally, 12 different pENTR/U6 shRNA target subclones, including target sequences to Lamin AC and Luciferase, were also recombined into these two DEST vectors. In all cases, the LxR crosses were efficient. When 2 $\mu$ l/20 $\mu$ l LxR reaction were transformed and 1/6th (50 $\mu$ l) of the transformation reaction plated, 300-800 colonies/plate were obtained in Top10 cells. Even in HB101 cells that were ~40 fold less competent to take up DNA than the Top10 cells, 10-20 colonies/plate could be obtained by plating more of the transformation reaction (100 $\mu$ l vs. 50 $\mu$ l). Note that the number of clones obtained are similar between the Lenti DEST and the Adeno DEST vectors, even though the



Adenoviral vector is almost 4 times the size of the Lentiviral vector (~36kb vs. ~8.6kb).

[0276] The LxR crosses were not only efficient but also effective. Ten out of ten of the Adeno DEST vector recombinants had the correct RNAi cassette as determined by RFLP analysis. pLenti DEST recombinants were transformed into both Top10 and HB101 *E. coli* cells because HB101 cells are known for reducing the recombination between the lentiviral LTR sequences. In this case, 10/10 recombinants were correct using HB101 cells.

#### shRNA Target Site Selection

[0277] The present invention may be used to create shRNAs with any desired stem length, orientation, and loop sequence. In general, target sequences should be complex (no runs of more than 3 of the same nucleotide), with low GC content (30-50%), and avoid known RNA-protein interaction sites. Target sites should be a minimum of 19nt, and sites of up to 29nt are effective.

#### DNA Oligo Insert Design

[0278] Once a candidate target site has been selected, it must be converted into an shRNA sequence, and the DNA oligos ordered for cloning into pENTR/U6. The shRNA sequence can be in two possible orientations. Either the sense target site or the antisense sequence of the target site can begin the shRNA, followed by a short loop sequence and then the opposite strand of the target site.

[0279] The fact that the polymerase (pol III) will terminate transcription after 4 thymidines (Ts) constrains the oligo design. Strings of more than 3 Ts should be avoided in the middle of a target site, or with any Ts in the connecting "loop", to prevent early termination. Additionally, Ts at the 3' end of the target will abut the polyT terminator and may cause slightly premature termination. Changing the sense/antisense orientation of the shRNA may be necessary for specific target sites to avoid early pol III termination by positioning different sequences next to the loop or polyT terminator.

[0280] Additionally, the native U6 snRNA initiates at a guanosine (G), and this +1 base is believed to be important. Although this system allows advanced users to choose any +1 base, we have designed all of our inserts to initiate at a G. In cases where the G is part of the target sequence, it is simply incorporated into the stem, with a complementary cytosine base placed just before the terminator. When G is not the first base in the sense or antisense target sequence, it is added to the 5' end of the shRNA with no complementary base at the 3' end. If use of a G is not desired, an A is believed to be better than an C or T.

[0281] Functional loops of anywhere from 4 to 11nt have been reported in the literature. Short loops are preferred as they reduce the lengths of the oligos needed for cloning. 5'-TTCG, 5'-AACG, and 5'CGAA have been used as the loop sequences in successful RNAi constructs. However, loops containing thymidines must be avoided in some cases as they may cause early termination, as discussed above.

[0282] Finally, to convert an shRNA sequence into an oligo pair for insertion, 5'CACC-3' was added to the 5' end of the shRNA sequence to create the "top" oligo. The "bottom" oligo is the complimentary sequence of the top oligo with the 5'CACC-3' removed and 5'AAAA-3' appended to the 5' end.

#### Conclusion

[0283] The pENTR/U6 and Gateway DEST vectors are the cornerstones of a superior system to clone shRNA target sequences into an RNAi expression cassette and deliver it to cells (Figure 28). Two other commercial sources with similar pol III vectors (Ambion with pSilencer, and OligoEngines with pSuper) require the synthesis of longer insert oligos (~70nt and 55nt respectively) because their cloning schemes need the end of the U6 promoter and termination signals to be "built-back" with the insert. Additionally, their cloning protocols call for ligation incubations of 1hr or greater compared to the 5 min bench-top reaction described here. This is likely due to the PEG present in the present ligation buffer, as well as the present vector design

features that eliminate background (the *ccdB* negative selection and the non-compatible ends left after *BsaI* digestion). The present invention also has the Gateway Advantage; any insert cloned and sequence verified in pENTR/U6 is then available for any application made possible by the DEST vectors – such as viral delivery of shRNA by Virapower™.

[0284] The demonstrations of RNAi in transient transfections reported here, as well as examples of successful RNAi by transduction indicate the U6 promoter can generate sufficient shRNA for RNAi. Experiments that define the rules required for efficient RNAi will make this vector all the more valuable.

## Example 2

### Expression of Interfering RNA using a Seamless Cloning Vector

#### Abstract and Introduction

[0285] Short hairpin RNA (shRNA) expression cassettes built into the U6 RNAi Entry Vector can be used to transiently knockdown genes of interest in cell culture. However, the Entry Vector carries no marker for selection in mammalian cells, and the plasmids must be introduced into cells by transfection. Transfection efficiency varies widely between cell lines and is ineffective in primary and terminally differentiated cells. In contrast to plasmid transfection, lentiviral delivery allows simple, stable transduction of a wide variety of cell types including primary and terminally differentiated cells. A number of recent publications describe the use of lentiviruses to deliver shRNAs to mammalian cells (Abbas-Terki *et al.* 2002, Dirac & Bernards 2003, Matta *et al.* 2003, Qin *et al.* 2003, Robinson *et al.* 2003, Stewart *et al.* 2003, Tiscornia *et al.* 2003), demonstrating an existing interest in this technique.

[0286] Invitrogen offers several Gateway-adapted lentiviral vectors for cloning of coding sequences downstream of a Pol II promoter. However, the presence of such an upstream promoter may interfere with Pol III expression from a U6 cassette. A promoterless Destination vector, pLenti6/RNAi-DEST has been created with *attR1* and *attR2* sites compatible with the U6 RNAi Entry Vector. A map of pLenti6/RNAi-DEST is shown in Figure 6A. pLenti6/RNAi-DEST allows simple and reliable transfer of shRNA expression cassettes into the lentiviral backbone. The viral vector confers blasticidin resistance for selection of stably transduced cells. Transduction by lentiviruses expressing lamin A/C shRNAs is demonstrated to efficiently and specifically knock down endogenous protein levels. pLenti6/RNAi-DEST complements the ViraPower™ product line and provides a powerful new application for the U6 RNAi Entry Vector.

[0287] Key Performance Criteria for Lenti6/RNAi-DEST include: (1) pLenti6/RNAi-DEST passing standard manufacturing QC specs for Destination vectors. (2) Gateway cloning shRNAs into pLenti6/RNAi-DEST and packaging virus at levels comparable with regular vectors. (3) Showing specific knockdown of endogenous lamin A/C gene.

#### Materials and Methods

[0288] Construction of pLenti6/RNAi-DEST Vector Lenti6/RNAi-DEST is the product of a Gateway BxP reaction between pLenti6/PL/attB4/V5/GW-GFP and pDONR 221. The BxP reaction was transformed into DB3.1 and selected on LB media containing Ampicillin (100 µg/ml) and chloramphenicol (15 µg/ml). Colonies of the transformants were analyzed by restriction digest. A map of pLenti6/RNAi DEST is shown in Figure 6A.

#### ShRNA-containing Entry Cones

[0289] The various shRNA-containing Entry clones used are set out in Table 1. The hairpins are targeted to sites on the lamin A/C or luciferase

genes as indicated. All entry clones were created by oligo cloning into pENTR/U6.2. Loops and stems choices are described in Example 1.

Table 1. pENTR/U6 Entry Clones

Clone name	Target gene	Orientation <sup>a</sup>	Loop sequence	Stem length <sup>b</sup> (bp)	Target position <sup>c</sup> (nt)
pENTR/U6-lamAC-SA-uucg	lamin A/C	SA	UUCG	19	610-628
pENTR/U6-lamAC-AS-uucg	lamin A/C	AS	UUCG	19	610-628
pENTR/U6-lamAC-AS-cgaa	lamin A/C	AS	CGAA	19	610-628
pENTR/U6-lamAC-SA-cgaa	lamin A/C	SA	CGAA	19	610-628
pENTR/U6-GL2-22	luciferase	AS	UUCG	22	153-174
pENTR/U6-GL2sh2 <sup>d</sup>	luciferase	AS	GAACGT TG	29	1355-1383

<sup>a</sup>Orientations are either sense-loop-antisense (SA) or antisense-loop-sense (AS).

<sup>b</sup>Stem length does not include +1 G base if it is not also part of the target site.

<sup>c</sup>Target position is relative to start codon.

<sup>d</sup>Hairpin design based on a previously assessed technology from Cold Spring Harbor Laboratories.

#### Destination Vector QC and generation of expression control vector

[0290] pLenti6/RNAi-DEST was monitored for quality using the official "Dest Vector QC Procedure" established by manufacturing. The expression control plasmid, pLenti6/RNAi/U6-GW/lamAC was generated by a standard Gateway LxR reaction between pLenti6/RNAi-DEST and pENTR/U6-lamAC-AS-cgaa. Clones of pLenti6/RNAi/U6-GW/lamAC were confirmed by restriction analyses. A map of pLenti6/RNAi/U6-GW/lamAC is shown in Figure 6B.

#### Cell culture

[0291] 293FT cells were cultured in DMEM/10% FBS/L-glutamine/non-essential amino acids/penicillin/streptomycin containing 500 µg/ml G418.

HeLa cells were cultured in DMEM/10% FBS/L-glutamine/non-essential amino acids/penicillin/streptomycin.

#### Virus production

[0292] For virus production,  $1 \times 10^7$  293FT cells were plated per T175 flask. Twenty-four hours later, culture medium was replaced with 20 ml OptiMem/10%FBS, and shRNA-encoding viruses were packaged by co-transfecting the 293FT cells with the respective lentiviral vector and pLP1, pLP2 and pLP/VSVG (at a mass ratio of 1:1:1:1, 24  $\mu$ g of total DNA) as follows: The 24  $\mu$ g DNA was mixed with 3 ml of OptiMem media. In a separate tube, 72  $\mu$ l of Lipofectamine 2000 was also mixed with 3 ml of OptiMem media. After a 5-minute incubation period at room temperature, the two mixtures were combined and incubated at room temperature for an additional 20 minutes. At the completion of the incubation period, the transfection mixture was added to the cells dropwise and the flask was gently rocked to mix. The following day the transfection complex was replaced with 30 ml complete media (DMEM, 10% FBS, 1% penicillin/streptomycin, L-glutamine and non-essential amino acids). Virus-containing media were harvested at day 2 and day 3 post-transfection, centrifuged at 3000 rpm for 5 minutes to remove dead cells, and filtered through sterile 0.45 micron cellulose acetate filters to remove fine debris. Viruses in the filtrates were concentrated by ultracentrifugation (90 minutes, 23000xg, 4°C). Viral pellets from ultra-centrifugation were resuspended in 500-600  $\mu$ l growth media. One hundred-microliter aliquots of concentrated virus were stored in -80°C freezer until use.

#### Viral Titering and Transduction

[0293] All applications of virus to cells were performed in the presence of 6  $\mu$ g/ml polybrene (Sigma, hexadimethrin bromide, #H9268) and media changes were performed 12-24 hours post transduction. For titering virus, 6-well

plates were seeded with  $2 \times 10^5$  HT1080 cells per well the day before transduction. One milliliter each of ten-fold serial dilutions of viral supernatant ranging from  $10^{-2}$  to  $10^{-8}$  was prepared. All dilutions were mixed by gentle inversion prior to adding to cells. Mock-transduced cells had no virus added to them. Plates were gently swirled to mix. The following day, the media was replaced with complete media. Forty-eight hours post-transduction, the cells were placed under  $10 \mu\text{g/ml}$  blasticidin selection. After 7 to 10 days of blasticidin selection the resulting colonies were stained with crystal violet : A 1% crystal violet solution was prepared in 10% ethanol. Each well was washed with 2 ml PBS followed by 1 ml of crystal violet solution for 10 minutes at room temperature. Excess stain was removed by two 2 ml PBS washes and colonies visible to the naked eye were counted to determine the viral titer of the original supernatants.

[0294] Transductions to test shRNA activities were performed in the appropriate cells in 12-well plates. Cells were plated at  $1 \times 10^5$ /well twenty-four hours before transduction. The next day, the media was replaced with complete media. Transduction was conducted in a final volume of 500  $\mu\text{l}$  and contained the appropriate volumes of virus supernatant to achieve a range of MOIs.

#### Cell lysis and Western Blot

[0295] Cell lysis for lamin A/C and beta-actin western blots were performed as follows: Forty-eight or 120 hours post-transduction, cells were harvested with Versene (Invitrogen), transferred to microfuge tubes, and centrifuged at 3000 RPM for 4min. Pellets were lysed in 2X NuPAGE® LDS Sample Buffer with 1X Sample Reducing Agent and denatured at  $95^\circ\text{C}$  for 5 min prior to electrophoresis. Protein samples were electrophoresed on NuPAGE® Novex 4-12% Tris-Bis Gels in 1X MOPS-SDS buffer with NuPAGE® Antioxidant in the upper chamber. Western blot analyses were performed using the Western Breeze Immunodetection Kit (Invitrogen) according to the manufacturer's protocol. Lamin A/C and beta-actin proteins were detected

using 1:1000 monoclonal anti-lamin A/C (BD Biosciences) and 1:5000 monoclonal anti-beta-actin (Abcam) antibodies, respectively.

## Results and Discussion

[0296] Destination Vector QC pLenti6/RNAi-DEST passed the standard manufacturing QC specs for Destination vectors with respect to total colony count (Table 2) and *ccdB* assay (Table 3).

### Virus Titers

[0297] ShRNA-encoding lentiviral vectors were used to produce virus in 293FT cells. The vectors produced viral titers comparable to titers attained with regular lentiviral vectors that do not contain shRNA (Table 4). This indicated that introduction of shRNAs into the lentiviral backbone does not compromise virus packaging or transduction efficiency.

Table 4. Lenti6/RNAi Virus Titers

Virus	Crude Virus Titer (cfu/ml)	Concentrated Virus Titer (cfu/ml) <sup>a</sup>
Lenti6/RNAi/U6-GW/lamAC-SA-uucg	1.00E+6	4.30E+08
Lenti6/RNAi/U6.2-GW/lamAC-AS-uucg	2.10E+6	5.85E+08
Lenti6/RNAi/U6.2-GW/amAC-AS-cgaa	8.00E+5	1.35E+08
Lenti6/RNAi/U6.2-GW/lamAC-SA-cgaa	1.20E+6	4.45E+08
Lenti6/RNAi/U6-GW/GL2-22	6.00E+5	4.50E+08
Lenti6/RNAi/U6-GW/GL2sh2	1.30E+6	5.20E+08
Lenti6/V5-GW/GFP(non-RNAi virus)	4.00E+5	8. 0E+07

<sup>a</sup>Concentrated from two 175cm<sup>2</sup> flasks each.

### Knockdown of Lamin A/C

[0298] Lentiviruses were tested for their ability to deliver shRNAs to specifically knock down lamin A/C expression in HeLa cells. Lentiviruses expressing luciferase-targeted shRNAs served as negative controls. Inhibition of lamin A/C expression was analyzed by western blot. ShRNAs targeted to lamin inhibited expression of both lamin A and C isoforms 48hr and 5 days



post-transduction (Figure 7). The extent of inhibition depended on transduced MOI, indicating knockdown was dose-dependent. Lentiviruses encoding shRNAs lamAC-AS-cgaa and lamAC-SA-cgaa provided the best lamin knockdowns (Figure 7, top panel lanes 11-16; bottom panel lanes 14-19). Of the two shRNAs, lamAC-AS-cgaa mediated robust inhibition even at the relatively low MOI of 14 (Figure 7, top panel lane 11 and bottom panel lane 14). The lamin A/C shRNAs had no effect on beta-actin expression irrespective of transduced MOI (Figure 7, beta-actin blots). Control luciferase shRNAs had no effect on beta-actin expression (Figure 7, top panel lanes 7-9 and 17-19; bottom panel lanes 1-3 and 11-13) and minor effect on lamin A/C expression even at the very high MOI of 520 (Figure 7, top panel lane 19; bottom panel lane 13). These results show specific inhibition of lamin expression with lamin-targeted shRNAs. The inhibition is not the effect of general inhibition of gene expression. Results of the control shRNA transduction provide further evidence of the specific activity of the lamin-directed shRNAs.

[0299] pLenti6/RNAi has also been used to specifically knock down luciferase (75% inhibition, 48 hrs post-transduction in Flp-In 293 luc cell line; data not shown) and lacZ at high MOIs (55% inhibition, 96 hrs post-transduction in HT1080LacZ cells; data not shown). These provide further evidence that pLenti6/RNAi-DEST vector will function with other RNAi cassettes.

## Summary

[0300] Gateway-adapted lentiviral vector pLenti6/RNAi-DEST has been developed for RNAi analyses. pLenti6/RNAi-DEST is designed to be used in LxR reactions with pENTR/U6. pLenti6/RNAi-DEST meets the performance criteria for all DEST vectors as well as criteria for packaging and transducing lentiviruses. Viruses Lenti6/RNAi/U6-GW/lamAC-AS-cgaa and Lenti6/RNAi/U6-GW/lamAC-SA-cgaa transduce shRNAs that specifically

knock down lamin A/C expression. The lamAC-AS-cgaa hairpin was chosen as the positive control for the U6 RNAi Entry and pLenti6/RNAi Kits. The sequence of lamAC-AS-cgaa hairpin is shown in the Kit Components and Configuration below.

### Example 3

#### RNAi using BLOCK-iT™ Dicer Kit

[0301] BLOCK-iT™ Kits (Invitrogen Corporation; Carlsbad, CA) can be used for fast and efficient RNAi applications. Eukaryotic cells naturally regulate gene expression with dsRNA. A BLOCK-iT™ Dicer Kit can be used to generate dsRNA that are then diced into siRNA, purified and transfected into cells. The BLOCK-iT™ Dicer Kit requires no expensive synthetic siRNAs. It also produces a pool of many siRNAs per gene, not just one or a few, which means a higher probability of knockdown (Figure 21,22, and 23). A purification procedure gives a high yield of siRNAs in a transfection-ready buffer and virtually eliminates remaining long dsRNA and cleave intermediates.

[0302] BLOCK-iT™ Long RNAi Transcription Kits use a T7 TOPO linker which allows any polymerase chain reaction (PCR) product to become a template for transcription (Figure 20). This mediates RNAi in invertebrates (e.g., insects, nematodes and protozoans), some mammalian embryonic cells (undifferentiated ES cells) and many mammalian cell lines after treatment with Dicer/RNase III. BLOCK-iT™ Kits allows for an inexpensive alternative to siRNA oligos. Exemplary uses of BLOCK-iT™ Kits are summarized in Figure 24.

Kit Components and Configurations

Complete Lentiviral RNAi Kit:

Components of the U6 RNAi Entry Vector Kit:

[0303] Purified, BsaI-linearized pENTR/U6.2; Annealed lamin A/C control  
oligos: Top 5'-CACCGTGTCTTCTGGAAGTCCAGCGAACT

GGACTTCCAGAAGAACA (SEQ ID NO:9), Bottom 5'-  
AAAATGTTCTTCTGGA

AGTCCAGTTCGCTGGACTTCCAGAAGAACAC (SEQ ID NO:10);

Sequencing primers: U6 forward 5'-GGACTATCATATGCTTACCG (SEQ ID NO:11), M13 reverse 5'-CAGGAAACAGCTATGAC (SEQ ID NO:12)(Catalog No. N530-02, Invitrogen Corp., Carlsbad, CA); T4 DNA ligase (Catalog No. 15224-025, Invitrogen Corp., Carlsbad, CA); 5X T4 DNA ligase buffer (Catalog No. Y90001, Invitrogen Corp., Carlsbad, CA Y90001); OneShot Top10 cells (Catalog No. C4040-03, Invitrogen Corp., Carlsbad, CA); pLenti6/RNAi/DEST; pLenti6/RNAi/U6-GW/lamAC; OneShot STBL3 cells; Virapower Bsd Lentiviral Support Kit (Catalog No. K4970-00, Invitrogen Corp., Carlsbad, CA); Gateway LR Clonase enzyme mix (Catalog No. 11791-091, Invitrogen Corp., Carlsbad, CA).

#### Lentiviral RNAi DEST Kit

[0304] pLenti6/RNAi/DEST; pLenti6/RNAi/U6-GW/lamAC; OneShot STBL3 cells; Gateway LR Clonase enzyme mix (Catalog No. 11791-019, Invitrogen Corp., Carlsbad, CA)

#### References:

Abbas-Terki *et al.*, Lentiviral-mediated RNA interference. *Hum Gene Ther.* 13:2197-2201 (2002)

Dirac & Bernards, Reversal of senescence in mouse fibroblasts through lentiviral suppression of p53. *J. Biol. Chem.* 278:11731-11734 (2003)

Matta *et al.*, Use of lentiviral vectors for delivery of small interfering RNA. *Cancer Biol. Ther.* 2:206-210 (2003)

Qin *et al.*, Inhibiting HIV-1 infection in human T cells by lentiviral-mediated delivery of small interfering RNA against CCR5. *Proc. Natl. Acad. Sci. (USA)* 100:183-188 (2003)

Rubinson *et al.*, A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. *Nat. Genet.* 33:401-406 (2003)

Stewart *et al.*, Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA*. 9:493-501 (2003)

Tiscornia *et al.*, A general method for gene knockdown in mice by using lentiviral vectors expressing small interfering RNA. *Proc. Natl. Acad. Sci. (USA)* 100:1844-1848 (2003)

Table 2. L x R Assay

Sample	Criteria	Values	Pass/Fail
Cells only	0 cfu/μg DNA	0 cfu/μg DNA	Pass
No DNA	0 cfu/μg DNA	0 cfu/μg DNA	Pass
DEST vector only	< 1100 cfu/μg DNA	660 cfu/μg DNA	Pass
LxR Reaction (n = 2)	$\geq 1.65 \times 10^6$ cfu/μg DNA	$2.31 \times 10^6$ cfu/μg DNA	Pass
pUC19 only (n = 2)	$\geq 7.5 \times 10^8$ cfu/μg DNA	$2.53 \times 10^{10}$ cfu/μg DNA	Pass

Table 3. ccdB Assay

Sample	Cell Type	Antibiotic	Transformation Efficiency
Cells Only	DB3.1	Amp	0 cfu/μg DNA
		Kan	0 cfu/μg DNA
pUC19 only (n=4)	DB3.1	Amp	$7.0 \times 10^6$ cfu/μg DNA
DEST vector only (n=4)	DB3.1	Amp	$3.0 \times 10^6$ cfu/μg DNA
Cells Only	TOP10	Amp	0 cfu/μg DNA
		Kan	0 cfu/μg DNA
pUC19 only (n=4)	TOP10	Amp	$2.65 \times 10^8$ cfu/μg DNA
DEST vector only (n=4)	TOP10	Amp	$5.75 \times 10^3$ cfu/μg DNA
		Kan	0 cfu/μg DNA
Fold-killing (criteria = $1 \times 10^4$ )			$2 \times 10^4$ Pass

[0305] The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations which is

not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising," "consisting essentially of," and "consisting of" may be replaced with either of the other two terms. The terms and expressions that have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed herein, optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the appended claims. In addition, where features or aspects of the invention are described in terms of Markush groups, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group.

[0306] The invention has been described broadly and generically herein. Each of the narrower species and subgeneric groupings falling within the generic disclosure also form part of the invention. This includes the generic description of the invention with a proviso or negative limitation removing any subject matter from the genus, regardless of whether or not the excised material is specifically recited herein. Other aspects of the invention are within the following claims.

[0307] All publications, patents and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference.

Table 5: pENTRU6 Vector Nucleic Acid Sequence

CTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCT  
TTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCA  
GCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAC  
CGCCTCTCCCCGCGCGTGGCCGATTCATTAATGCAGCTGGCACGA  
CAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAA  
TACGCGTACCGCTAGCCAGGAAGAGTTTGTAGAAACGCAAAAAGG  
CCATCCGTCAGGATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTA  
TGGCGGGCGTCCTGCCCCGCCACCCTCCGGGCCGTTGCTTCACAACG  
TTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCA  
CCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAG  
CCTTTCGTTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCGTTAAC  
GCTAGCATGGATGTTTTCCAGTCACGACGTTGTAAAACGACGGCC  
AGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGAAGTATAGT  
GACCTGTTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGC  
CAACTTTGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGA  
CTGGATCCGGTACCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCAT  
GATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATA  
ATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATA  
CGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAA  
ATTATGTTTTTAAATGGACTATCATATGCTTACCGTAACTTGAAAGT  
ATTTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACC  
GGAGACCGCGGCCGCTGGATCCGGCTTACTAAAAGCCAGATAACA  
GTATGCGTATTTGCGCGCTGATTTTTGCGGTATAAGAATATATACTG  
ATATGTATACCCGAAGTATGTCAAAAAGAGGTGTGCTATGAAGCA  
GCGTATTACAGTGACAGTTGACAGCGACAGCTATCAGTTGCTCAAG  
GCATATATGATGTCAATATCTCCGGTCTGGTAAGCACAACCATGCA  
GAATGAAGCCCGTCGTCTGCGTGCCGAACGCTGGAAAGCGGAAAA  
TCAGGAAGGGATGGCTGAGGTCGCCCCGGTTTATTGAAATGAACGG  
CTCTTTTGCTGACGAGAACAGGGACTGGTGAAATGCAGTTTAAGGT  
TTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTA  
CAGAGTGATATTATTGACACGCCCGGGCGACGGATGGTGATCCCCC  
TGGCCAGTGACAGTCTGCTGTCAGATAAAGTCTCCCGTGAACCTTA  
CCCGGTGGTGCATATCGGGGATGAAAGCTGGCGCATGATGACCAC  
CGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAAGTGGCT  
GATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACCTG  
ATGTTCTGGGGAATATAAGGTCTCATTTTTTTTCTAGACCCAGCTTT  
CTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTT  
GCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCA  
TCCAGCTGATATCCCCTATAGTGAGTCGTATTACATGGTCATAGCT  
GTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTACAT  
TGCACAAGATAAAAATATATCATCATGAACAATAAACTGTCTGCT  
TACATAAACAGTAA

Table 5 (continued): pENTRU6 Vector Nucleic Acid Sequence

TACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCC  
GCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGG  
GCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGCTTGT  
ATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAG  
GTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTG  
GCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTA  
CTCCTGATGATGCATGGTACTCACCCTGCGATCCCCGGAAAAAC  
AGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATT  
GTTGATGCGCTGGCAGTGTTCCCTGCGCCGGTTGCATTCGATTCCTGT  
TTGTAATTGTCCTTTTAAACAGCGATCGCGTATTTTCGTCTCGCTCAGG  
CGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTGAT  
GACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATG  
CATAAACTTTTGCCATTCTCACC GGATT CAGTCGTCACTCATGGTGA  
TTTCTCACTTGATAACCTTATTTT GACGAGGGGAAATTAATAGGTT  
GTATTGATGTTGGACGAGTCGGAATCGCAGACCGATAACCAGGATCT  
TGCCATCCTATGGAAC TGCCTCGGTGAGTTTTCTCCTTCATTACAGA  
AACGGCTTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAA  
ATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGAATTGG  
TTAATTGGTTGTAACTGGCAGAGCATTACGCTGACTTGACGGGA  
CGGCGCAAGCTCATGACCAAAATCCCTTAACGTGAGTTACGCGTCG  
TTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT  
GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAA  
ACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCA  
ACTCTTTTTCCGAAGGTAAC TGGCTTCAGCAGAGCGCAGATACCAA  
ATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAA  
CTCTGTAGCACCGCCTACATACTCGCTCTGCTAATCCTGTTACCAG  
TGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTC  
AAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGG  
GGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCG  
AACTGAGATACCTACAGCGTGAGCATTGAGAAAGCGCCACGCTTCC  
CGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGG  
AACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTA  
TCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGAT  
TTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAG  
CAACGCGGCCTTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTC  
ACATGTT SEQ ID NO:1

Table 6: Nucleotide sequence of plasmid pLenti6/V5-DEST.

AATGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACG  
ATGAGTTAGCAACATGCCTTACAAGGAGAGAGAAAAAGCACCGTGCA  
TGCCGATTGGTGGAAAGTAAGGTGGTACGATCGTGCCTTATTAGGAA  
GGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTGC  
CGCATTGCAGAGATATTGTATTTAAGTGCCTAGCTCGATACATAAA  
CGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGC  
TAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAG  
TGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAG  
AGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTG  
GCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGCT  
CTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGG  
CGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGG  
AGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCG  
GGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAGG  
GGGAAAGAAAAAATATAAATTAAACATATAGTATGGGCAAGCAG  
GGAGCTAGAACGATTTCGCAGTTAATCCTGGCCTGTTAGAAACATCA  
GAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAG  
ACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCC  
TCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAG  
CTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCG  
CACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATATG  
AGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAA  
ATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTG  
GTGCAGAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTTCCTT  
GGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATG  
ACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGC  
AGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTT  
GCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCT  
GGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTG  
GGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAAT  
GCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGGAATCACACGA  
CCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAA  
TACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATG  
AACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTG  
GTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATG  
ATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGTGTACTTTC  
TATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAG  
ACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATA  
GAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTTCGATTA  
GTGAACGGATCTCGACGGTATCGATAAGCTTGGGAGTTCCGCGTTA



Table 6 (continued).

Nucleotide sequence of plasmid pLenti6/V5-DEST.

CATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCC  
CCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCA  
ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAA  
CTGCCCACCTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCC  
CCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCC  
CAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGT  
ATTAGTCATCGCTATTACCATGGTGATGCGGTTTTTGGCAGTACATC  
AATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC  
ACCCCATGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAACG  
GGACTTTCCAAAATGTCGTAACAACCTCCGCCCCATTGACGCAAATG  
GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTT  
TAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGA  
CCTCCATAGAAGACACCGACTCTAGAGGATCCACTAGTCCAGTGTG  
GTGGAATTCTGCAGATATCAACAAGTTTGTACAAAAAAGCTGAACG  
AGAAACGTAAAATGATATAAATATCAATATATTAAATTAGATTTTG  
CATAAAAAACAGACTACATAATACTGTAAAACACAACATATCCAG  
TCACTATGGCGGCCCGCATTAGGCACCCCAGGCTTTACACTTTATGC  
TTCCGGCTCGTATAATGTGTGGATTTTGAGTTAGGATCCGGCGAGA  
TTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGA  
TATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTG  
AGGCATTTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCA  
GCTGGATATTACGGCCTTTTTTAAAGACCGTAAAGAAAAATAAGCAC  
AAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGC  
TCATCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATA  
TGGGATAGTGTTACACCCTTGTTACACCGTTTTTCCATGAGCAAACCTG  
AAACGTTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCA  
GTTTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAAC  
CTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTC  
AGCCAATCCCTGGGTGAGTTTCACCAAGTTTTGATTAAACGTGGCC  
AATATGGACAACCTTCTTCGCCCCCGTTTTTCACCATGGGCAAATATT  
ATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTTCAGGTTCA  
TCATGCCGTCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAA  
TTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAAAGATCT  
GGATCCGGCTTACTAAAAGCCAGATAACAGTATGCGTATTTGCGCG  
CTGATTTTTTGCGGTATAAGAATATATACTGATATGTATACCCGAAG  
TATGTCAAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACA  
GTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAA  
TATCTCCGGTCTGGTAAGCACAACCATGCAGAATGAAGCCCGTCGT  
CTGCGTGCCGAACGCTGGAAAGCGGAAAATCAGGAAGGGATGGCT  
GAGGTCGCCCCGGTTTATTGAAATGAACG

Table 6 (continued). Nucleotide sequence of plasmid pLenti6/V5-DEST.

GCTCTTTTGCTGACGAGAACAGGGACTGGTGAAATGCAGTTTAAGG  
TTTACACCTATAAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGT  
ACAGAGTGATATTATTGACACGCCCGGGCGACGGATGGTGATCCCC  
CTGGCCAGTGCACGTCTGCTGTCAGATAAAGTCTCCCGTGAACCTT  
ACCCGGTGGTGCATATCGGGGATGAAAGCTGGCGCATGATGACCA  
CCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAAGTGGC  
TGATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACCTG  
ATGTTCTGGGGAATATAAATGTCAGGCTCCGTTATACACAGCCAGT  
CTGCAGGTCGACCATAGTGACTGGATATGTTGTGTTTTACAGTATT  
ATGTAGTCTGTTTTTTATGCAAAATCTAATTTAATATATTGATATTT  
ATATCATTTTACGTTTTCTCGTTCAGCTTTCTTGTACAAAGTGGTTGA  
TATCCAGCACAGTGGCGGCCGCTCGAGTCTAGAGGGCCCGCGGTTTC  
GAAGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACGC  
GTACCGGTTAGTAATGAGTTTGGAATTAATTCTGTGGAATGTGTGT  
CAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGGCAGGCAGAAG  
TATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAG  
TCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCA  
ATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCC  
CCTAACTCCGCCCAGTTCCGCCCATCTCCGCCCATGGCTGACTAA  
TTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTA  
TTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCA  
AAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAGCA  
CGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAAT  
ACGACAAGGTGAGGAACTAAACCATGGCCAAGCCTTTGTCTCAAG  
AAGAATCCACCCTCATTGAAAGAGCAACGGCTACAATCAACAGCA  
TCCCCATCTCTGAAGACTACAGCGTCGCCAGCGCAGCTCTCTCTAG  
CGACGGCCGCATCTTCACTGGTGTCAATGTATATCATTTTACTGGG  
GGACCTTGTGCAGAACTCGTGGTGCTGGGCACTGCTGCTGCTGCGG  
CAGCTGGCAACCTGACTTGTATCGTCGCGATCGGAAATGAGAACAG  
GGGCATCTTGAGCCCCTGCGGACGGTGCCGACAGGTGCTTCTCGAT  
CTGCATCCTGGGATCAAAGCCATAGTGAAGGACAGTGATGGACAG  
CCGACGGCAGTTGGGATTCGTGAATTGCTGCCCTCTGGTTATGTGT  
GGGAGGGCTAAGCACAAATTCGAGCTCGGTACCTTTAAGACCAATG  
ACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGG  
GGGGACTGGAAGGGCTAATCACTCCCAACGAAGACAAGATCTGC  
TTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGG  
GAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAA  
GCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGAC  
TCT

Table 6 (continued). Nucleotide sequence of plasmid pLenti6/V5-DEST.

GGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCT  
CTAGCAGTAGTAGTTCATGTCATCTTATTATTCAGTATTTATAACTT  
GCAAAGAAATGAATATCAGAGAGTGAGAGGAACTTGTTTATTGCA  
GCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAA  
ATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTC  
ATCAATGTATCTTATCATGTCTGGCTCTAGCTATCCCGCCCCTAACT  
CCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCC  
CCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCT  
CGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGG  
CCTAGGGACGTACCCAATTCGCCCTATAGTGAGTCGTATTACGCGC  
GCTCACTGGCCGTCGTTTTTACAACGTCGTGACTGGGAAAACCCTGG  
CGTTACCCAACCTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCT  
GGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTT  
GCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATT  
AAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTT  
GCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCCTTTCTC  
GCCACGTTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCC  
CTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCCAAAAA  
ACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAG  
ACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGG  
ACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATT  
CTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAAA  
AATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATAT  
TAACGCTTACAATTTAGGTGGCACTTTTCGGGGGAAATGTGCGCGGA  
ACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCT  
CATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGG  
AAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTT  
TGCGGCATTTTGCCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGA  
AAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACA  
TCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCC  
CGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGT  
GGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTC  
GCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGT  
CACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATG  
CAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTT  
CTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACA  
ACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCT  
GAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGT  
AGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTT  
ACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATA  
AAGTTGCAGGACCACCTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTT

Table 6 (continued). Nucleotide sequence of plasmid pLenti6/V5-DEST.

TATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATC  
ATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTA  
TCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGAC  
AGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTC  
AGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATT  
TTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATG  
ACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACC  
CCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGC  
GTAATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGG  
TTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAC  
GGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGC  
CGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATA  
CCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT  
AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATA  
AGGCGCAGCGGTCTGGGCTGAACGGGGGGTTCGTGCACACAGCCCA  
GCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTG  
AGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACA  
GGTATCCGGTAAGCGGCAGGGTCTGGAACAGGAGAGCGCACGAGGG  
AGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTT  
TCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGG  
GGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTT  
CCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATC  
CCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGAT  
ACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGC  
GAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCG  
CGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACT  
GGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCA  
CTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGGCTCGTATG  
TTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCT  
ATGACCATGATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGA  
ACAAAAGCTGGAGCTGCAAGCTT SEQ ID NO:2

Table 7. Nucleotide sequence of plasmid pLenti6/V5-dTOPO™.

AATGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACG  
ATGAGTTAGCAACATGCCTTACAAGGAGAGAAAAAGCACCGTGCA  
TGCCGATTGGTGGGAAGTAAGGTGGTACGATCGTGCCTTATTAGGAA  
GGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTGC  
CGCATTGCAGAGATATTGTATTTAAGTGCCTAGCTCGATACATAAA  
CGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGC  
TAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAG  
TGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAG  
AGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTG  
GCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGCT  
CTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGG  
CGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGA CTAGCGG  
AGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCG  
GGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAGG  
GGGAAAGAAAAAATATAAATTAACATATAGTATGGGCAAGCAG  
GGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCA  
GAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAG  
ACAGGATCAGAAGA ACTTAGATCATTATATAATACAGTAGCAACCC  
TCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAG  
CTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCG  
CACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATATG  
AGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAA  
ATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTG  
GTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTT  
GGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATG  
ACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGC  
AGCAGAACAAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTT  
GCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCT  
GGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTG  
GGGTTGCTCTGGAAA ACTCATTTGCACCACTGCTGTGCCTTGGAAT  
GCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGGAATCACACGA  
CCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAA  
TACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATG  
AACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTG  
GTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATG  
ATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTC  
TATAGTGAATAGAGTTAGGCAGGGATATTCACCATTCGTTTCAG  
ACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATA  
GAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTA  
GTGAACGGATCTCGACGGTATCGATAAGCTTGGGAGTTCCGCGTTA

Table 7 (continued). Nucleotide sequence of plasmid pLenti6/V5-dTOPO™.

CATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCC  
CCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCA  
ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAA  
CTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCC  
CCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCC  
CAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGT  
ATTAGTCATCGCTATTACCATGGTGATGCGGTTTTTGGCAGTACATC  
AATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC  
ACCCCAATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG  
GGACTTTCCAATAATGTCGTAACAACCTCCGCCCCATTGACGCAAATG  
GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTT  
TAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGA  
CCTCCATAGAAGACACCGACTCTAGAGGATCCACTAGTCCAGTGTG  
GTGGAATTGATCCCTTCACCAAGGGCTCGAGTCTAGAGGGCCCGCG  
GTTCGAAGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTA  
CGCGTACCGGTTAGTAATGAGTTTGGAATTAATTCTGTGGAATGTG  
TGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGGCAGGCAG  
AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGA  
AAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCAT  
CTCAATTAGTCAGCAACCATAGTCCCGCCCCCTAACTCCGCCCCATCC  
CGCCCCCTAACTCCGCCCAGTTCCGCCCCATTCTCCGCCCCATGGCTGA  
CTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCTGCCTCTGA  
GCTATTCCAGAAGTAGTGAGGAGGCTTTTTTTGGAGGCCCTAGGCTTT  
TGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATC  
AGCACGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTA  
TAATACGACAAGGTGAGGAACTAAACCATGGCCAAGCCTTTGTCTC  
AAGAAGAATCCACCCTCATTGAAAGAGCAACGGCTACAATCAACA  
GCATCCCCATCTCTGAAGACTACAGCGTCGCCAGCGCAGCTCTCTC  
TAGCGACGGCCGCATCTTCACTGGTGTCAATGTATATCATTTTACTG  
GGGGACCTTGTGCAGAACTCGTGGTGCTGGGCACTGCTGCTGCTGC  
GGCAGCTGGCAACCTGACTTGTATCGTCGCGATCGGAAATGAGAAC  
AGGGGCATCTTGAGCCCCTGCGGACGGTGCCGACAGGTGCTTCTCG  
ATCTGCATCCTGGGATCAAAGCCATAGTGAAGGACAGTGATGGAC  
AGCCGACGGCAGTTGGGATTCGTGAATTGCTGCCCTCTGGTTATGT  
GTGGGAGGGCTAAGCACAATTCGAGCTCGGTACCTTTAAGACCAAT  
GACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTTAAAAGAAAAG  
GGGGGACTGGAAGGGCTAATCACTCCCAACGAAGACAAGATCTG  
CTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTG  
G

Table 7 (continued). Nucleotide sequence of plasmid pLenti6/V5-dTOPO™.

GAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAA  
GCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCCGTCTGTTGTGTGAC  
TCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAA  
TCTCTAGCAGTAGTAGTTCATGTCATCTTATTATTCAGTATTTATAA  
CTTGCAAAGAAATGAATATCAGAGAGTGAGAGGAACTTGTTTATTG  
CAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTTAC  
AAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAAC  
TCATCAATGTATCTTATCATGTCTGGCTCTAGCTATCCCGCCCCCTAA  
CTCCGCCCATCCCGCCCCCTAACTCCGCCCAGTTCCGCCCATCTCCG  
CCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGC  
CTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGG  
GGCCTAGGGACGTACCCAATTCGCCCTATAGTGAGTCGTATTACGC  
GCGCTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCT  
GGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCA  
GCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACA  
GTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGC  
ATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA  
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCCTCCCTTCCTTT  
CTCGCCACGTTTCGCCGGCTTTCGCCGTCAAGCTCTAAATCGGGGGC  
TCCCTTTAGGGTTCGGATTTAGTGCTTTACGGCACCTCGACCCCAA  
AACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGAT  
AGACGGTTTTTCGCCCTTTCGCGTTGGAGTCCACGTTCTTTAATAGT  
GGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCT  
ATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTA  
AAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAA  
TATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGGAAATGTGCGC  
GGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCC  
GCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAA  
AGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTT  
TTTTGCGGCATTTTGCCCTTCCTGTTTTTGCTCACCCAGAAACGCTGG  
TGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTT  
ACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCG  
CCCCGAAGAACGTTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTA  
TGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCG  
GTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACC  
AGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATT  
ATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTA  
CTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGC  
ACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGA  
GCTGAATGAAGCCAT



Table 7 (continued). Nucleotide sequence of plasmid pLenti6/V5-dTOPO™.

ACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAAC  
AACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTAGCTTCC  
CGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGA  
CCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAA  
ATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTG  
GGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGG  
GGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGA  
TAGGTGCCTCACTGATTAAGCATTGGTAACCTGTCAGACCAAGTTTA  
CTCATATATACTTTAGATTGATTTAAACTTCATTTTAAATTTAAAA  
GGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCC  
TTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAG  
ATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTG  
CTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCG  
GATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCA  
GAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGG  
CCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTG  
CTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTC  
TTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGC  
GGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGC  
GAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAG  
AAAGCGCCACGCTTCCCGAAGGGGAGAAAGGCGGACAGGTATCCGG  
TAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAG  
GGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTCGCCACCT  
CTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGC  
CTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCT  
TTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATT  
CTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCG  
CCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGC  
GGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCG  
ATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGG  
GCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGC  
ACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAA  
TTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGA  
TTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCT  
GGAGCTGCAAGCTT SEQ ID NO:3



Table 8. Nucleotide sequence of pLenti4/V5-DEST.

AATGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACG  
ATGAGTTAGCAACATGCCTTACAAGGAGAGAAAAAGCACCGTGCA  
TGCCGATTGGTGGGAAGTAAGGTGGTACGATCGTGCCTTATTAGGAA  
GGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTGC  
CGCATTGCAGAGATATTGTATTTAAGTGCCTAGCTCGATACATAAA  
CGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGC  
TAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAG  
TGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAG  
AGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTG  
GCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGCT  
CTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGG  
CGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGG  
AGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCG  
GGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAGG  
GGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAG  
GGAGCTAGAACGATTTCGCAGTTAATCCTGGCCTGTTAGAAACATCA  
GAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAG  
ACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCC  
TCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAG  
CTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCG  
CACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATATG  
AGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAA  
ATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTG  
GTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTTCCTT  
GGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATG  
ACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGC  
AGCAGAACAAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTT  
GCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCT  
GGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTG  
GGGTTGCTCTGGAAAACCTATTTGCACCACTGCTGTGCCTTGGAAT  
GCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGGAATCACACGA  
CCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAA  
TACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATG  
AACAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTG  
GTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATG  
ATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTC  
TATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAG  
ACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATA  
GAAGAAGAAGGTGGAGAGAGA

Table 8 (continued). Nucleotide sequence of pLenti4/V5-DEST.

GACAGAGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATC  
GATAAGCTTGGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCC  
GCCTGGCTGACCGCCCAACGACCCCGCCATTGACGTCAATAATG  
ACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTC  
AATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCA  
AGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGT  
AAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACT  
TTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG  
GTGATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTG  
ACTCACGGGGGATTTCCAAGTCTCCACCCCATGACGTCAATGGGAG  
TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAAC  
AACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGG  
GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCT  
GGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGACT  
CTAGAGGATCCACTAGTCCAGTGTGGTGGAATTCTGCAGATATCAA  
CAAGTTTGTACAAAAAAGCTGAACGAGAAACGTAAAATGATATAA  
ATATCAATATATTAAATTAGATTTTGCATAAAAAACAGACTACATA  
ATACTGTAAACACAACATATCCAGTCACTATGGCGGCCCGCATTAG  
GCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGG  
ATTTTGAGTTAGGATCCGGCGAGATTTTCAGGAGCTAAGGAAGCTA  
AAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCA  
ATGGCATCGTAAAGAACATTTTGAGGCATTTTCAGTCAGTTGCTCAA  
TGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAA  
AGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCA  
CATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCCGTATGGCA  
ATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTACCCCTTGTT  
ACACCGTTTTTCATGAGCAAACCTGAAACGTTTTTCATCGCTCTGGAG  
TGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAA  
GATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGT  
TTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTC  
ACCAGTTTTGATTTAAACGTGGCCAATATGGACAACCTTCTTCGCCC  
CCGTTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCT  
GATGCCGCTGGCGATTTCAGGTTTCATCATGCCGTCTGTGATGGCTTC  
CATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGT  
GGCAGGGCGGGGCGTAAAGATCTGGATCCGGCTTACTAAAAGCCA  
GATAACAGTATGCGTATTTGCGCGCTGATTTTTGCGGTATAAGAAT  
ATATACTGATATGTATACCCGAAG

Table 8 (continued). Nucleotide sequence of pLenti4/V5-DEST.

TATGTCAAAAAGAGGGTGTGCTATGAAGCAGCGTATTACAGTGACA  
GTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAA  
TATCTCCGGTCTGGTAAGCACAAACCATGCAGAATGAAGCCCGTCGT  
CTGCGTGCCGAACGCTGGAAAGCGGAAAATCAGGAAGGGATGGCT  
GAGGTCGCCCCGGTTTATTGAAATGAACGGCTCTTTTGCTGACGAGA  
ACAGGGACTGGTGAAATGCAGTTTAAGGTTTACACCTATAAAAGA  
GAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTG  
ACACGCCCGGGCGACGGATGGTGATCCCCCTGGCCAGTGCACGTCT  
GCTGTCAGATAAAGTCTCCCGTGAACCTTTACCCGGTGGTGCATATC  
GGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCCAGTGTG  
CCGGTCTCCGTTATCGGGGAAGAAGTGGCTGATCTCAGCCACCGCG  
AAAATGACATCAAAAACGCCATTAACCTGATGTTCTGGGGGAATATA  
AATGTCAGGCTCCGTTATACACAGCCAGTCTGCAGGTCGACCATAG  
TGACTGGATATGTTGTGTTTTACAGTATTATGTAGTCTGTTTTTTAT  
GCAAAATCTAATTTAATATATTGATATTTATATCATTTTACGTTTCT  
CGTTCAGCTTTTCTTGTACAAAGTGGTTGATATCCAGCACAGTGGCG  
GCCGCTCGAGTCTAGAGGGCCCGCGGTTCTGAAGGTAAGCCTATCCC  
TAACCCTCTCCTCGGTCTCGATTCTACGCGTACCGGTTAGTAATGAG  
TTTGGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAG  
TCCCCAGGCTCCCCAGGCAGGCAGAAAGTATGCAAAGCATGCATCTC  
AATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCA  
GGCAGAAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATA  
GTCCCGCCCCCTAACTCCGCCCATCCCGCCCCCTAACTCCGCCCAGTTC  
CGCCCATTTCTCCGCCCCATGGCTGACTAATTTTTTTTTTATTTATGCAG  
AGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGG  
AGGCTTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCCCTGTTG  
ACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACAA  
GGTGAGGAACTAAACCATGGCCAAGTTGACCAGTGCCGTTCCGGTG  
CTCACCGCGCGCGACGTGCGCGGAGCGGTCTGAGTTCTGGACCGACC  
GGCTCGGGTTCTCCCGGGACTTCGTGGAGGACGACTTCGCCGGTGT  
GGTCCGGGACGACGTGACCCTGTTTCATCAGCGCGGTCCAGGACCAG  
GTGGTGCCGGACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGG  
ACGAGCTGTACGCCGAGTGGTCGGAGGTCGTGTCCACGAACTTCCG  
GGACGCCTCCGGGGCCGGCCATGACCGAGATCGGCGAGCAGCCGTG  
GGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGTGCA  
CTTCGTGGCCGAGGAGCAGGACTGACACGTGCTACGAGATTTAAAT  
GGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGC  
CACTTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATCACTCCC  
AACGAAGACAAGATCTGCTTTTTTGCTTGTACTGGGTCTCTCTGGTTA  
GACCAGATCTGAGCCTGGGAGCTCTCTG

Table 8 (continued). Nucleotide sequence of pLenti4/V5-DEST.

GCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTG  
AGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACT  
AGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAG  
TAGTAGTTCATGTCATCTTATTATTCAGTATTTATAACTTGCAAAGA  
AATGAATATCAGAGAGTGAGAGGAACTTGTTTATTGCAGCTTATAA  
TGGTTACAAATAAAGCAATAGCATCACAAATTCACAAATAAAGC  
ATTTTTTTCCTGCACTTCTAGTTGTGGTTTGTCCAAACTCATCAATG  
TATCTTATCATGTCTGGCTCTAGCTATCCCGCCCCCTAACTCCGCCCA  
TCCCGCCCCCTAACTCCGCCCAGTTCCGCCCATCTCCGCCCCATGGC  
TGAATAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTC  
TGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGG  
ACGTACCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACT  
GGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACC  
CAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTA  
ATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAG  
CCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGC  
GGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGC  
GCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCCTTTCTCGCCACG  
TTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGG  
GTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGAT  
TAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTT  
TTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTG  
TTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGA  
TTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAAAAATGAG  
CTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGC  
TTACAATTTAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCT  
ATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAG  
ACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAG  
TATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTTTGCGG  
CATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTA  
AAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAA  
CTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAG  
AACGTTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGC  
GGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCCGC  
ATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAG  
AAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTG  
CTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGAC  
AACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATG  
GGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATG  
AAGCCATACCAAACGAC

Table 8 (continued). Nucleotide sequence of pLenti4/V5-DEST.

GAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGC  
AAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGGCAACAAT  
TAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGC  
GCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGC  
CGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGAT  
GGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG  
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCT  
CACTGATTAAGCATTGGTAAGTGTGACACCAAGTTTACTCATATAT  
ACTTTAGATTGATTTAAACTTCATTTTAAATTTAAAGGATCTAGG  
TGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGA  
GTTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGA  
TCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAAC  
AAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAG  
CTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGA  
TACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTC  
AAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTT  
ACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTG  
GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCTGGGCTGA  
ACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTAC  
ACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACG  
CTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG  
GTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCC  
TGGTATCTTTATAGTCCTGTCTGGGTTTCGCCACCTCTGACTTGAGCG  
TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAC  
GCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTT  
TGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACC  
GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC  
GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCC  
AATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGC  
AGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGC  
AACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTT  
TACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGA  
TAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGC  
GCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTGCAAG  
CTT SEQ ID NO:4

Table 9. Nucleotide sequence of pLenti6/UbC/V5-DEST.

AATGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACG  
ATGAGTTAGCAACATGCCTTACAAGGAGAGAGAAAAAGCACCGTGCA  
TGCCGATTGGTGGGAAGTAAGGTGGTACGATCGTGCCTTATTAGGAA  
GGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTGC  
CGCATTGCAGAGATATTGTATTTAAGTGCCTAGCTCGATACATAAA  
CGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGC  
TAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAG  
TGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAG  
AGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTG  
GCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGCT  
CTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGG  
CGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGG  
AGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCG  
GGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAGG  
GGGAAAGAAAAAATATAAATTAACATATAGTATGGGCAAGCAG  
GGAGCTAGAACGATTTCGCAGTTAATCCTGGCCTGTTAGAAACATCA  
GAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAG  
ACAGGATCAGAAGAACTTAGATCATTATATAACAGTAGCAACCC  
TCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAG  
CTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCG  
CACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATATG  
AGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAA  
ATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTG  
GTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTTCCTT  
GGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATG  
ACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGC  
AGCAGAACAAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTT  
GCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCT  
GGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTG  
GGGTTGCTCTGGAAAACCTATTGCAACCACTGCTGTGCCTTGGAAT  
GCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGGAATCACACGA  
CCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAA  
TACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATG  
AACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTG  
GTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATG  
ATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTC  
TATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAG  
ACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATA  
GAAGAAGAAGGTGGAGAGAGA

Table 9 (continued). Nucleotide sequence of pLenti6/UbC/V5-DEST.

GACAGAGACAGATCCATTCGATTAGTGAAACGGATCTCGACGGTATC  
GGATCTGGCCTCCGCGCCGGGTTTTGGCGCCTCCCGCGGGCGCCCC  
CCTCCTCACGGCGAGCGCTGCCACGTCAGACGAAGGGCGCAGGAG  
CGTCCTGATCCTTCCGCCCCGACGCTCAGGACAGCGGCCCGCTGCT  
CATAAGACTCGGCCTTAGAACCCCAAGTATCAGCAGAAGGACATTTT  
AGGACGGGACTTGGGTGACTCTAGGGCACTGGTTTTCTTTCCAGAG  
AGCGGAACAGGCGAGGAAAAGTAGTCCCTTCTCGGCGATTCTGCG  
GAGGGATCTCCGTGGGGCGGTGAACGCCGATGATTATATAAGGAC  
GCGCCGGGTGTGGCACAGCTAGTTCCGTGCGAGCCGGGATTTGGGT  
CGCGGTCTTGTGTTGTGGATCGCTGTGATCGTCACTTGGTGAGTAGC  
GGGCTGCTGGGCTGGCCGGGGCTTTCGTGGCCGCCGGGGCCGCTCGG  
TGGGACGGAAGCGTGTGGAGAGACCGCCAAGGGCTGTAGTCTGGG  
TCCGCGAGCAAGGTTGCCCTGAACTGGGGGTTGGGGGGAGCGCAG  
CAAAATGGCGGCTGTTCCCGAGTCTTGAATGGAAGACGCTTGTGAG  
GCGGGCTGTGAGGTCGTTGAAACAAGGTGGGGGGCATGGTGGGCG  
GCAAGAACCCAAGGTCTTGAGGCCTTCGCTAATGCGGGAAAGCTCT  
TATTCGGGTGAGATGGGCTGGGGCACCATCTGGGGACCCTGACGTG  
AAGTTTGTCACTGACTGGAGAACTCGGTTTGTGCTCTGTTGCGGGG  
GCGGCAGTTATGCGGTGCCGTTGGGCAGTGCACCCGTACCTTTGGG  
AGCGCGCGCCCTCGTCGTGTCGTGACGTCACCCGTTCTGTTGGCTTA  
TAATGCAGGGTGGGGCCACCTGCCGGTAGGTGTGCGGTAGGCTTTT  
CTCCGTCGCAGGACGCAGGGTTCGGGCCTAGGGTAGGCTCTCCTGA  
ATCGACAGGCGCCGGACCTCTGGTGAGGGGAGGGATAAGTGAGGC  
GTCAGTTTCTTTGGTCGGTTTTATGTACCTATCTTCTTAAGTAGCTG  
AAGCTCCGGTTTTTGAACATATGCGCTCGGGGTTGGCGAGTGTGTTTT  
GTGAAGTTTTTTAGGCACCTTTTTGAAATGTAATCATTTGGGTCAATA  
TGTAATTTTTCAGTGTTAGACTAGTAAATTGTCCGCTAAATTCTGGCC  
GTTTTTGGCTTTTTTTGTTAGACGAAGCTTGGTACCGAGCTCGGATCC  
ACTAGTCCAGTGTGGTGGAATTCTGCAGATATCAACAAGTTTGTAC  
AAAAAAGCTGAACGAGAAACGTAAAATGATATAAATATCAATATA  
TTAAATTAGATTTTGCATAAAAAACAGACTACATAATACTGTAAAA  
CACAACATATCCAGTCACTATGGCGGCCGCATTAGGCACCCCAGGC  
TTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGATTTTGAGTTA  
GGATCCGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAA  
AAAAATCACTGGATATAACACCGTTGATATATCCCAATGGCATCGT  
AAAGAACATTTTGAAGGCATTTTCAGTCAGTTGCTCAATGTACCTATA  
ACCAGACCGTTCAGCTGGATATTACGGCCTTTTTTAAAGACCGTAAA  
GAAAAATAAGCACAAAGTTTTATCCGGCCTTTATTCACATTCTTGCCC  
GCC



Table 9 (continued). Nucleotide sequence of pLenti6/UbC/V5-DEST.

TGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAAGACGGTGA  
GCTGGTGATATGGGATAGTGTTACCCCTTGTTACACCGTTTTCCATG  
AGCAAACCTGAAACGTTTTTCATCGCTCTGGAGTGAATACCACGACGA  
TTTCCGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTTACG  
GTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTT  
TTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAAGTTTTGATTAA  
ACGTGGCCAATATGGACAACCTTCTTCGCCCCCGTTTTTCACCATGGG  
CAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATT  
CAGGTTTCATCATGCCGTCTGTGATGGCTTCCATGTCGGCAGAATGC  
TTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGT  
AAAGATCTGGATCCGGCTTACTAAAAGCCAGATAACAGTATGCGTA  
TTTGCGCGCTGATTTTTTGCGGTATAAGAATATATACTGATATGTATA  
CCCGAAGTATGTCAAAAAGAGGTGTGCTATGAAGCAGCGTATTAC  
AGTGACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATG  
ATGTCAATATCTCCGGTCTGGTAAGCACAAACCATGCAGAATGAAGC  
CCGTCGTCTGCGTGCCGAACGCTGGAAAGCGGAAAATCAGGAAGG  
GATGGCTGAGGTCGCCCCGGTTTATTGAAATGAACGGCTCTTTTGCT  
GACGAGAACAGGGACTGGTGAAATGCAGTTTAAGGTTTACACCTAT  
AAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATA  
TTATTGACACGCCCGGGCGACGGATGGTGATCCCCCTGGCCAGTGC  
ACGTCTGCTGTCAGATAAAGTCTCCCGTGAACTTTACCCGGTGGTG  
CATATCGGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCC  
AGTGTGCCGGTCTCCGTTATCGGGGAAGAAGTGGCTGATCTCAGCC  
ACCGCGAAAATGACATCAAAAACGCCATTAACCTGATGTTCTGGGG  
AATATAAATGTCAGGCTCCGTTATACACAGCCAGTCTGCAGGTCGA  
CCATAGTGAAGTGGATATGTTGTGTTTTACAGTATTATGTAGTCTGTT  
TTTTATGCAAAATCTAATTTAATATATTGATATTTATATCATTTTAC  
GTTTCTCGTTCAGCTTTCTTGTAACAAGTGGTTGATATCCAGCACAG  
TGGCGGGCCGCTCGAGTCTAGAGGGGCCGCGGTTTCAAGGTAAGCCT  
ATCCCTAACCCCTCTCCTCGGTCTCGATTCTACGCGTACCGGTTAGTA  
ATGAGTTTGGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTG  
GAAAGTCCCCAGGCTCCCCAGGCAGGCAGAAGTATGCAAAGCATG  
CATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCC  
CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAA  
CCATAGTCCCGCCCCCTAACTCCGCCCCATCCCGCCCCCTAACTCCGCCC  
AGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTTATTTA  
TGCAGAGGCCGAGGCCGCCT



Table 9 (continued). Nucleotide sequence of pLenti6/UbC/V5-DEST.

CTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGG  
CCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCG  
GATCTGATCAGCACGTGTTGACAATTAATCATCGGCATAGTATATC  
GGCATAGTATAATACGACAAGGTGAGGAACTAAACCATGGCCAAG  
CCTTTGTCTCAAGAAGAATCCACCCTCATTGAAAGAGCAACGGCTA  
CAATCAACAGCATCCCCATCTCTGAAGACTACAGCGTCGCCAGCGC  
AGCTCTCTCTAGCGACGGCCGCATCTTCACTGGTGTCAATGTATATC  
ATTTTACTGGGGGACCTTGTGCAGAACTCGTGGTGTCTGGGCACTGC  
TGCTGCTGCGGCAGCTGGCAACCTGACTTGTATCGTCGCGATCGGA  
AATGAGAACAGGGGCATCTTGAGCCCCTGCGGACGGTGCCGACAG  
GTGCTTCTCGATCTGCATCCTGGGATCAAAGCCATAGTGAAGGACA  
GTGATGGACAGCCGACGGCAGTTGGGATTCGTGAATTGCTGCCCTC  
TGGTTATGTGTGGGAGGGCTAAGCACAATTCGAGCTCGGTACCTTT  
AAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTAA  
AAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAACGAAGA  
CAAGATCTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGAT  
CTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAG  
CCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCT  
GTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCA  
GTGTGGAAAATCTCTAGCAGTAGTAGTTCATGTCATCTTATTATTCA  
GTATTTATAACTTGCAAAGAAATGAATATCAGAGAGTGAGAGGAA  
CTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATC  
ACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGG  
TTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGCTCTAGCTAT  
CCCGCCCCCTAACTCCGCCCCTCCCGCCCCCTAACTCCGCCCAGTTCCG  
CCCATTCTCCGCCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAG  
GCCGAGGCCGCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAG  
GCTTTTTTGGAGGCCTAGGGACGTACCCAATTCGCCCTATAGTGAG  
TCGTATTACGCGCGCTCACTGGCCGTCGTTTTACAACGTCGTGACTG  
GGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGACGACATCCC  
CCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCC  
CTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTG  
TAGCGGCGCATTAAGCGCGGGCGGGTGTGGTGGTTACGCGCAGCGT  
GACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCT  
TCCCTTCCTTTCTCGCCACGTTCCGCCGGCTTTCCCGTCA

Table 9 (continued). Nucleotide sequence of pLenti6/UbC/V5-DEST.

AGCTCTAAATCGGGGGCTCCCTTTAGGGTTCGATTTAGTGCTTTAC  
GGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAG  
TGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGT  
CCACGTTCTTTAATAGTGGACTCTTGTTCCAACTGGAACAACACT  
CAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGA  
TTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTAA  
CGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTT  
TTCGGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATA  
CATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGC  
TTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCG  
TGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGC  
TCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTT  
GGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAA  
GATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGC  
ACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGC  
CGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGAC  
TTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCA  
TGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAA  
CACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGA  
GCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCTT  
GATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAG  
CGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAA  
CTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAA  
TAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC  
GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGT  
GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTA  
AGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAAC  
TATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACT  
GATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTT  
AGATTGATTTAAAACTTCATTTTAAATTTAAAAGGATCTAGGTGAA  
GATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTT  
CGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTC  
TTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAA  
AACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACC  
AACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCA  
AATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGA  
ACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCA  
GTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACT  
CAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGGCTGAACGG  
GGGGTTCGTGCACACAGCCAG

Table 9 (continued). Nucleotide sequence of pLenti6/UbC/V5-DEST.

CTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGA  
GCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAG  
GTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGA  
GCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTT  
CGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGG  
GGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTT  
CCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATC  
CCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGAT  
ACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGC  
GAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCG  
CGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACT  
GGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCA  
CTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGGCTCGTATG  
TTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCT  
ATGACCATGATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGA  
ACAAAAGCTGGAGCTGCAAGCTT SEQ ID NO:5

Table 10. Nucleotide sequence of plasmid pLP1.

TTGGCCCATTCGCATACGTTGTATCCATATCATAATATGTACATTTAT  
ATTGGCTCATGTCCAACATTACCGCCATGTTGACATTGATTATTGAC  
TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCA  
TATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTG  
GCTGACCGCCCAACGACCCCCGCCCATGACGTCAATAATGACGTA  
TGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGG  
GTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGT  
ATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATG  
GCCCCGCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCT  
ACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGAT  
GCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCA  
CGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT  
TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACCTC  
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGT  
CTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGA  
CGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGAT  
CCAGCCTCCCCTCGAAGCTTACATGTGGTACCGAGCTCGGATCCTG  
AGAACTTCAGGGTGAGTCTATGGGACCCTTGATGTTTTCTTTCCCCT  
TCTTTTCTATGGTTAAGTTCATGTCATAGGAAGGGGAGAAGTAACA  
GGGTACACATATTGACCAAATCAGGGTAATTTTGCATTTGTAATTTT  
AAAAAATGCTTTCTTCTTTTAATACTTTTTTGTATTCTTATTTCT  
AATACTTTCCCTAATCTCTTTCTTTCAGGGCAATAATGATACAATGT  
ATCATGCCTCTTTGCACCATTTCTAAAGAATAACAGTGATAATTTCTG  
GGTTAAGGCAATAGCAATATTTCTGCATATAAATATTTCTGCATAT  
AAATTGTAAGTATGTAAGAGGTTTCATATTGCTAATAGCAGCTAC  
AATCCAGCTACCATTCTGCTTTTATTTTATGGTTGGGATAAGGCTGG  
ATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTTCATACC  
TCTTATCTTCCTCCCACAGCTCCTGGGCAACGTGCTGGTCTGTGTGC  
TGGCCCATCACTTTGGCAAAGCACGTGAGATCTGAATTCGAGATCT  
GCCGCCGCATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAA  
TTAGATCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGGAAAGAAA  
AAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAA  
CGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTA  
GACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAG  
AAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGT  
GCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAA  
GATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAAGCACAGCAAG  
CAGCAGCTGACACAGGACACAGCAATCAGGTCAGCCAAAATTACC  
CTATAGTGCAGAACATCCAGGGGGCAAATGGTACATCAGGCCATATC  
ACCTAGAACTTTAAATGCATGGG

Table 10 (continued). Nucleotide sequence of plasmid pLP1

TAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCA  
TGTTTTTCAGCATTATCAGAAGGAGCCACCCACAAAGATTTAACAC  
CATGCTAAACACAGTGGGGGGACATCAAGCAGCCATGCAAATGTT  
AAAAGAGACCATCAATGAGGAAGCTGCAGAATGGGATAGAGTGCA  
TCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAACC  
AAGGGGAAGTGACATAGCAGGAAGTACTAGTACCCTTCAGGAACA  
AATAGGATGGATGACACATAATCCACCTATCCCAGTAGGAGAAAT  
CTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAAT  
GTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGA  
ACCCTTTAGAGACTATGTAGACCGATTCTATAAACTCTAAGAGCC  
GAGCAAGCTTCACAAGAGGTAAAAAATTGGATGACAGAAACCTTG  
TTGGTCCAAAATGCGAACCCAGATTGTAAGACTATTTTAAAAGCAT  
TGGGACCAGGAGCGACACTAGAAGAAATGATGACAGCATGTCAGG  
GAGTGGGGGGACCCGGCCATAAAGCAAGAGTTTTGGCTGAAGCAA  
TGAGCCAAGTAACAAATCCAGCTACCATAATGATACAGAAAGGCA  
ATTTTAGGAACCAAAGAAAGACTGTTAAGTGTTTCAATTGTGGCAA  
AGAAGGGCACATAGCCAAAAATTGCAGGGCCCCTAGGAAAAAGGG  
CTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTAC  
TGAGAGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCCCACAAG  
GGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCC  
CCACCAGAAGAGAGCTTCAGGTTTGGGGAAGAGACAACAACCTCCC  
TCTCAGAAGCAGGAGCCGATAGACAAGGAAGTGTATCCTTTAGCTT  
CCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATA  
GGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGAT  
ACAGTATTAGAAGAAATGAATTTGCCAGGAAGATGGAAACCAAAA  
ATGATAGGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATC  
AGATACTCATAGAAATCTGCGGACATAAAGCTATAGGTACAGTATT  
AGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACT  
CAGATTGGCTGCACTTTAAATTTTCCCATTAGTCCTATTGAGACTGT  
ACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAGTTAAACA  
ATGGCCATTGACAGAAGAAAAAATAAAAGCATTAGTAGAAATTTG  
TACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTGGGCCTGA  
AAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGT  
ACTAAATGGAGAAAATTAGTAGATTTTCAGAGAACTTAATAAGAGA  
ACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCACATCCTGCAG  
GGTTAAACAGAAAAAATCAGTAACAGTACTGGATGTGGGCGATG  
CATATTTTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATACTGC  
ATTTACCATACCTAGTATAACAATGAGACACCAGGGATTAGATAT  
CAGTACAATGTGCTTCCACAGGGA

Table 10 (continued). Nucleotide sequence of plasmid pLP1

TGGAAAGGATCACCAGCAATATTCCAGTGTAGCATGACAAAAATCT  
TAGAGCCTTTTAGAAAAACAAAATCCAGACATAGTCATCTATCAATA  
CATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCAT  
AGAACAAAAATAGAGGAACTGAGACAACATCTGTTGAGGTGGGGA  
TTTACCACACCAGACAAAAACATCAGAAAGAACCTCCATTTCCTTT  
GGATGGGTATGAACTCCATCCTGATAAATGGACAGTACAGCCTAT  
AGTGCTGCCAGAAAAGGACAGCTGGACTGTCAATGACATACAGAA  
ATTAGTGGGAAAATTGAATTGGGCAAGTCAGATTTATGCAGGGATT  
AAAGTAAGGCAATTATGTAACTTCTTAGGGGAACCAAAGCACTA  
ACAGAAGTAGTACCACTAACAGAAGAAGCAGAGCTAGAACTGGCA  
GAAAACAGGGAGATTCTAAAAGAACCGGTACATGGAGTGTATTAT  
GACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAA  
GGCCAATGGACATATCAAATTTATCAAGAGCCATTTAAAAATCTGA  
AAACAGGAAAGTATGCAAGAATGAAGGGTGCCCACTAATGATG  
TGAAACAATTAACAGAGGCAGTACAAAAAATAGCCACAGAAAGCA  
TAGTAATATGGGGAAAGACTCCTAAATTTAAATTACCCATACAAAA  
GGAAACATGGGAAGCATGGTGGACAGAGTATTGGCAAGCCACCTG  
GATTCCTGAGTGGGAGTTTGTCAATACCCCTCCCTTAGTGAAGTTAT  
GGTACCAGTTAGAGAAAGAACCCATAATAGGAGCAGAAACTTTCT  
ATGTAGATGGGGCAGCCAATAGGGGAACTAAATTAGGAAAAGCAG  
GATATGTAAGTACAGAGGAAGACAAAAAGTTGTCCCCCTAACGG  
ACACAACAAATCAGAAGACTGAGTTACAAGCAATTCATCTAGCTTT  
GCAGGATTCGGGATTAGAAGTAAACATAGTGACAGACTCACAATA  
TGCATTGGGAATCATTCAAGCACAACCAGATAAGAGTGAATCAGA  
GTTAGTCAGTCAAATAATAGAGCAGTTAATAAAAAAGGAAAAAGT  
CTACCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGA  
ACAAGTAGATAAATTGGTCAGTGCTGGAATCAGGAAAGTACTATTT  
TTAGATGGAATAGATAAGGCCCAAGAAGAACATGAGAAATATCAC  
AGTAATTGGAGAGCAATGGCTAGTGATTTTAACCTACCACCTGTAG  
TAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAG  
GGGAAGCCATGCATGGACAAGTAGACTGTAGCCCAGGAATATGGC  
AGCTAGATTGTACACATTTAGAAGGAAAAGTTATCTTGGTAGCAGT  
TCATGTAGCCAGTGGATATATAGAAGCAGAAGTAATTCCAGCAGA  
GACAGGGCAAGAAACAGCATACTTCCTCTTAAAATTAGCAGGAAG  
ATGGCCAGTAAAAACAGTACATACAGACAATGGCAGCAATTTTAC  
CAGTACTACAGTTAAGGCCGCCTGTTGGTGGGCGGGGATCAAGCA  
GGAATTTGGCATTCCCTACAATCCCCAAAGTCAAGGAGTAATAGAA

Table 10 (continued). Nucleotide sequence of plasmid pLP1

TCTATGAATAAAGAATTAAAGAAAATTATAGGACAGGTAAGAGAT  
CAGGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTATTCATCC  
ACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGG  
AAAGAATAGTAGACATAATAGCAACAGACATACAACTAAAGAAT  
TACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAG  
GGACAGCAGAGATCCAGTTTGGAAAGGACCAGCAAAGCTCCTCTG  
GAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATAAA  
AGTAGTGCCAAGAAGAAAAGCAAAGATCATCAGGGATTATGGAAA  
ACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGA  
TTAACACATGGAATTCCGGAGCGGCCGCAGGAGCTTTGTTCCCTTGG  
GTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATGAC  
GCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGACAGCAG  
CAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGC  
AACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGG  
CTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTGTTGGG  
GTTGCTCTGGAAAACATTTGCACCACTGCTGTGCCTTGGAATGCT  
AGTTGGAGTAATAAATCTCTGGAACAGATTTGGAATCACACGACCT  
GGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTCCGCG  
GAATTCACCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTG  
GTGTGGCTAATGCCCTGGCCCACAAGTATCACTAAGCTCGCTTTCTT  
GCTGTCCAATTTCTATTAAAGGTTCTTTGTTCCCTAAGTCCAATA  
CTAAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGC  
CTAATAAAAAACATTTATTTTCATTGCAATGATGTATTTAAATTATT  
TCTGAATATTTTACTAAAAAGGGAATGTGGGAGGTCAGTGCATTTA  
AAACATAAAGAAATGAAGAGCTAGTTCAAACCTTGGGAAAATACA  
CTATATCTTAAACTCCATGAAAGAAGGTGAGGCTGCAAACAGCTAA  
TGCACATTGGCAACAGCCCCTGATGCCTATGCCTTATTCATCCCTCA  
GAAAAGGATTCAAGTAGAGGCTTGATTTGGAGGTTAAAGTTTTGCT  
ATGCTGTATTTTACATTACTTATTGTTTTAGCTGTCCTCATGAATGT  
CTTTTCACTACCCATTTGCTTATCCTGCATCTCTCAGCCTTGACTCC  
ACTCAGTTCTCTTGCTTAGAGATACCACCTTTCCCCTGAAGTGTTCC  
TTCCATGTTTTACGGCGAGATGGTTTCTCCTCGCCTGGCCACTCAGC  
CTTAGTTGTCTCTGTTGTCTTATAGAGGTCTACTTGAAGAAGGAAA  
AACAGGGGGCATGGTTTGACTGTCCTGTGAGCCCTTCTTCCCTGCCT  
CCCCACTCACAGTGACCCGGAATCCCTCGACATGGCAGTCTAGCA  
CTAGTGCGGCCGCAGATCTGCTTCCTCGCTCACTGACTCGCTGCGCT  
CGGTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGT  
AATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACAT  
GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCG  
CGTTGCTGGCGTTTTTCCATAGGCTCCGCC



Table 10 (continued). Nucleotide sequence of plasmid pLP1

CCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC  
GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAA  
GCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATAC  
CTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTC  
ACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTG  
GGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTAT  
CCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC  
GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTA  
TGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGC  
TACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG  
TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAAC  
CACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACG  
CGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGG  
GGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGT  
CATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAA  
AAATGAAGTTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGT  
CTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGAT  
CTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA  
TAACTACGATACGGGAGGGGCTTACCATCTGGCCCCAGTGCTGCAAT  
GATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATA  
AACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACT  
TTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGT  
AAGTAGTTCGCCAGTTAATAGTTTTCGCGCAACGTTGTTGCCATTGCT  
ACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCA  
GCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTT  
GTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGA  
AGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGC  
ATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACT  
GGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGAC  
CGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATAACCGCGCCACA  
TAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG  
CGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGT  
AACCCACTCGTGACCCAACTGATCTTCAGCATCTTTTACTTTCACC  
AGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAA  
AAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCC  
TTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGC  
GGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTC  
CGCGCACATTTCCCCGAAAAGTGCCACCTGACGGGATCCCCTGAGG  
GGGCCCCCATGGGCTAGAGGATCCGGCCTCGGCCTCTGCATAAATA  
AAAAAAATTAGTCAGCCATGAGC SEQ ID NO:6



Table 11. Nucleotide sequence of plasmid pLP2.

AATGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACG  
ATGAGTTAGCAACATGCCTTACAAGGAGAGAAAAAGCACCGTGCA  
TGCCGATTGGTGGAAGTAAGGTGGTACGATCGTGCCTTATTAGGAA  
GGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTCC  
GCATTGCAGAGATATTGTATTTAAGTGCCTAGCTCGATACAATAAA  
CGCCATTTGACCATTACACCATTTGGTGTGCACCTCCAAGCTCGAG  
CTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTG  
TTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCCCTCG  
AAGCTAGTCGATTAGGCATCTCCTATGGCAGGAAGAAGCGGAGAC  
AGCGACGAAGACCTCCTCAAGGCAGTCAGACTCATCAAGTTTCTCT  
ATCAAAGCAACCCACCTCCCAATCCCGAGGGGACCCGACAGGCCC  
GAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATC  
CATTCGATTAGTGAACGGATCCTTAGCACTTATCTGGGACGATCTG  
CGGAGCCTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTACTCTT  
GATTGTAACGAGGATTGTGGAACCTTCTGGGACGCAGGGGGGTGGGA  
AGCCCTCAAATATTGGTGGAATCTCCTACAATATTGGAGTCAGGAG  
CTAAAGAATAGTGCTGTTAGCTTGCTCAATGCCACAGCTATAGCAG  
TAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTACAAGAAGCTT  
GGCACTGGCCGTCGTTTTACAACGTCGTGATCTGAGCCTGGGAGAT  
CTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTG  
CCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGG  
TAACTAGAGATCAGGAAAACCCTGGCGTTACCCAACTTAATCGCCT  
TGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCC  
CGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAAT  
GGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCA  
CACCGCATACGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGC  
ATTAAGCGCGGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA  
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTT  
CTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGC  
TCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAA  
AACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGAT  
AGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGT  
GGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGGCT  
ATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTA  
AAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAA  
TATTAACGTTTACAATTTTATGGTGCACCTCTCAGTACAATCTGCTCT  
GATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTG  
ACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACA  
A

Table 11 (continued). Nucleotide sequence of plasmid pLP2.

GCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTTCACCGT  
CATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTAT  
TTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGT  
GGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTT  
CTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGA  
TAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAAC  
ATTTCCGTGTCGCCCTTATTCCCTTTTTTTCGGGCATTTTGCCTTCCTG  
TTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGA  
TCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGC  
GGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTCCAATGA  
TGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATT  
GACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGA  
ATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGA  
TGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGT  
GATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGA  
AGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCG  
CCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGA  
CGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCG  
CAAACCTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAA  
TTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGC  
GCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGC  
CGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGAT  
GGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG  
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCT  
CACTGATTAAGCATTGGTAACCTGTCAGACCAAGTTTACTCATATAT  
ACTTTAGATTGATTTAAAACCTTCATTTTTTAATTTAAAAGGATCTAGG  
TGAAGATCCTTTTTTGATAATCTCATGACCAAATCCCTTAACGTGA  
GTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGA  
TCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAAC  
AAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAG  
CTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGA  
TACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTC  
AAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTT  
ACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTG  
GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTTCGGGCTGA  
ACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTAC  
ACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACG  
CTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG  
G

Table 11 (continued). Nucleotide sequence of plasmid pLP2.

TCGGAACAGGAGAGCGCACGAGGGGAGCTTCCAGGGGGGAAACGCCT  
GGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGT  
CGATTTTTGTGATGCTCGTCAGGGGGGGCGGAGCCTATGGAAAAACG  
CCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTT  
GCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCG  
TATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACG  
ACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCA  
ATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCA  
GCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCA  
ACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTT  
ACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGAT  
AACAATTTACACAGGAAACAGCTATGACATGATTACGAATTCGAT  
GTACGGGCCAGATATACGCGTATCTGAGGGGACTAGGGTGTGTTTA  
GGCGAAAAGCGGGGCTTCGGTTGTACGCGGTTAGGAGTCCCCTCAG  
GATATAGTAGTTTCGCTTTTGCATAGGGAGGGGGA SEQ ID NO:7

Table 12. Nucleotide sequence of plasmid pLP/VSVG.

TTGGCCCATTCGCATACGTTGTATCCATATCATAATATGTACATTTAT  
ATTGGCTCATGTCCAACATTACCGCCATGTTGACATTGATTATTGAC  
TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCA  
TATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTG  
GCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTA  
TGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGG  
GTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGT  
ATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATG  
GCCCCGCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCT  
ACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGAT  
GCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCA  
CGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT  
TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACCTC  
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGT  
CTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGA  
CGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGAT  
CCAGCCTCCCCTCGAAGCTTACATGTGGTACCGAGCTCGGATCCTG  
AGAACTTCAGGGTGAGTCTATGGGACCCTTGATGTTTTCTTTCCCCT  
TCTTTTCTATGGTTAAGTTCATGTCATAGGAAGGGGAGAAGTAACA  
GGGTACACATATTGACCAAATCAGGGTAATTTTGCATTTGTAATTTT  
AAAAAATGCTTTTCTTCTTTTAATACTTTTTTGTATTATCTTATTTCT  
AATACTTTCCCTAATCTCTTTCTTTCAGGGCAATAATGATACAATGT  
ATCATGCCTCTTTGCACCATTTCTAAAGAATAACAGTGATAATTTCTG  
GGTTAAGGCAATAGCAATATTTCTGCATATAAATATTTCTGCATAT  
AAATTGTAAGTATGTAAGAGGTTTCATATTGCTAATAGCAGCTAC  
AATCCAGCTACCATTTCTGCTTTTTATTTTATGGTTGGGATAAGGCTGG  
ATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTTCATACC  
TCTTATCTTCCTCCCACAGCTCCTGGGCAACGTGCTGGTCTGTGTGC  
TGGCCCATCACTTTGGCAAAGCACGTGAGATCTGAATTCTGACACT  
ATGAAGTGCCTTTTGTACTTAGCCTTTTTATTCAATTGGGGTGAATTG  
CAAGTTCACCATAGTTTTTCCACACAACCAAAAAGGAAACTGGAAA  
AATGTTCCCTTCTAATTACCATTATTGCCCGTCAAGCTCAGATTTAAA  
TTGGCATAATGACTTAATAGGCACAGCCTTACAAGTCAAAATGCCC  
AAGAGTCACAAGGCTATTCAAGCAGACGGTTGGATGTGTCATGCTT  
CCAAATGGGTCACTACTTGTGATTTCGCTGGTATGGACCGAAGTA  
TATAACACATTCCATCCGATCCTTCACTCCATCTGTAGAACAATGC  
AAGGAAAGCATTGAACAAACGAAACAAGGAACTTGGCTGAATCCA  
GGCTTCCCTCCTCAAAGTTGTGGATATGCAACTGTGACGGATGCCG

Table 12 (continued). Nucleotide sequence of plasmid pLP/VSVG.

AAGCAGTGATTGTCCAGGTGACTCCTCACCATGTGCTGGTTGATGA  
ATACACAGGAGAATGGGTTGATTCACAGTTCATCAACGGAAAATG  
CAGCAATTACATATGCCCCACTGTCCATAACTCTACAACCTGGCAT  
TCTGACTATAAGGTCAAAGGGCTATGTGATTCTAACCTCATTTCAT  
GGACATCACCTTCTTCTCAGAGGACGGAGAGCTATCATCCCTGGGA  
AAGGAGGGGCACAGGGTTCAGAAGTAACTACTTTGCTTATGAACTG  
GAGGCAAGGCCTGCAAAATGCAATACTGCAAGCATTGGGGAGTCA  
GACTCCCATCAGGTGTCTGGTTCGAGATGGCTGATAAGGATCTCTT  
TGCTGCAGCCAGATTCCCTGAATGCCCAGAAGGGTCAAGTATCTCT  
GCTCCATCTCAGACCTCAGTGGATGTAAGTCTAATTCAGGACGTTG  
AGAGGATCTTGGATTATTCCTCTGCCAAGAAACCTGGAGCAAAAT  
CAGAGCGGGTCTTCCAATCTCTCCAGTGGATCTCAGCTATCTTGCTC  
CTAAAAACCCAGGAACCGGTCCTGCTTTCACCATAATCAATGGTAC  
CCTAAAATACTTTGAGACCAGATACATCAGAGTCGATATTGCTGCT  
CCAATCCTCTCAAGAATGGTCGGAATGATCAGTGGAACCTACCACAG  
AAAGGGAACCTGTGGGATGACTGGGCACCATATGAAGACGTGGAAA  
TTGGACCCAATGGAGTTCTGAGGACCAGTTCAGGATATAAGTTTCC  
TTTATACATGATTGGACATGGTATGTTGGACTCCGATCTTCATCTTA  
GCTCAAAGGCTCAGGTGTTGGAACATCCTCACATTCAAGACGCTGC  
TTCGCAACTTCCTGATGATGAGAGTTTATTTTTTGGTGATACTGGGC  
TATCCAAAAATCCAATCGAGCTTGTAGAAGGTTGGTTCAGTAGTTG  
GAAAAGCTCTATTGCCTCTTTTTTCTTTATCATAGGGTTAATCATTG  
GACTATTCTTGGTTCTCCGAGTTGGTATCCATCTTTGCATTAAATTA  
AAGCACACCAAGAAAAGACAGATTTATACAGACATAGAGATGAAC  
CGACTTGGAAAGTAACTCAAATCCTGCACAACAGATTCTTCATGTT  
TGGACCAAATCAACTTGTGATACCATGCTCAAAGAGGCCTCAATTA  
TATTTGAGTTTTTAATTTTTATGAAAAAAAAAAAAAAAAACGGAAT  
TCACCCACACAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGT  
GGCTAATGCCCTGGCCACAAGTATCACTAAGCTCGCTTTCTTGCT  
GTCCAATTTCTATTAAAGGTTCTTTGTTCCCTAAGTCCAACCTACTA  
AACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTA  
ATAAAAAACATTTATTTTCATTGCAATGATGTATTAAATTATTTCT  
GAATATTTTACTAAAAAGGGAATGTGGGAGGTCAGTGCATTAAAAA  
CATAAAGAAATGAAGAGCTAGTTCAAACCTTGGGAAAATACACTA  
TATCTTAACTCCATGAAAGAAGGTGAGGCTGCAAACAGCTAATGC  
ACATTGGCAACAGCCCCTGATGCCTATGCCTTATTCATCCCTCAGA  
AAAGGATTCAAGTAGAGGCTTGATTTGGAGGTTAAAGTTTTGCTAT  
GCTGTATTTTACATTACTTATTGTTTTAGCTGTCCTCATGAATGTCTT  
TTCCTACCCATTTGCTTATCCTGCATCTCTCAGCCTTGACTCCACT  
CAGTTCTCTTGCTTAGAGATAACACCTTTCCCCTGAAGTGTTCTTC  
CATGTTTTACGGCGAGATGGTTTCTCCTCGCCT

Table 12 (continued). Nucleotide sequence of plasmid pLP/VSVG.

GGCCACTCAGCCTTAGTTGTCTCTGTTGTCTTATAGAGGTCTACTTG  
AAGAAGGAAAAACAGGGGGGCATGGTTTGA CTGTCCTGTGAGCCCT  
TCTTCCCTGCCTCCCCCACTCACAGTGACCCGGAATCCCTCGACATG  
GCAGTCTAGCACTAGTGCGGGCCGCAGATCTGCTTCCTCGCTCACTG  
ACTCGCTGCGCTCGGTCGTTTCGGCTGCGGGCAGCGGTATCAGCTCA  
CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC  
AGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC  
GTAAAAAGGCCGCGTTGCTGGCGTTTTTTCATAGGCTCCGCCCCC  
TGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAA  
CCCGACAGGACTATAAAGATAACAGGCGTTTCCCCCTGGAAGCTCC  
CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC  
CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCT  
GTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTG  
TGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGT  
AACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC  
TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAG  
GCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACAC  
TAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC  
TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCG  
CTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAG  
AAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCT  
GACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGA  
GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATG  
AAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGAC  
AGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTC  
TATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACT  
ACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAC  
CGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCA  
GCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATC  
CGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGT  
AGTTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTACAG  
GCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCC  
GGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCA  
AAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA  
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT  
TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGA  
GTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGT  
TGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCA  
GAACTTTAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAA  
ACTCTCAAGGATCTTACCGCTGTTGA

Table 12 (continued). Nucleotide sequence of plasmid pLP/VSVG.

GATCCAGTTCGATGTAACCCACTCGTGACCCCAACTGATCTTCAGC  
ATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGG  
CAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGA  
ATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGG  
TTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAAT  
AAACAAATAGGGGTTCGCGCACATTTCCCCGAAAAGTGCCACCTG  
ACGGGATCCCCTGAGGGGGCCCCCATGGGCTAGAGGATCCGGCCT  
CGGCCTCTGCATAAATAAAAAAAATTAGTCAGCCATGAGC SEQ ID  
NO:8

WHAT IS CLAIMED IS:

1. A method of producing an RNA molecule for use as an interfering RNA comprising:

- (a) identifying one or more target nucleic acid sequences;
- (b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein said interfering RNAs bind to said one or more target nucleic acid sequences;
- (c) combining
  - (i) one or more first nucleic acid molecules encoding one or more interfering RNAs flanked by one or more first type IIs restriction enzyme recognition sites;
  - (ii) one or more second nucleic acid molecules comprising one or more selectable markers flanked by one or more second type IIs restriction enzyme recognition sites; and
  - (iii) one or more site-specific type IIs restriction enzymes; and
- (d) incubating said combination under conditions sufficient to join one or more of said nucleic acid molecules encoding one or more interfering RNAs and one or more of said second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules;
- (e) inserting said one or more product nucleic acid molecules into a host cell; and
- (f) expressing said one or more interfering RNAs in said host cell.

2. The method of claim 1, wherein said first and second restriction sites are the same.

3. The method of claim 1, wherein said first and second restriction sites are different.



4. The method of claim 1, wherein said first or second nucleic acid molecule is a vector.

5. The method of claim 1, wherein said first or second nucleic acid molecule is a linear nucleic acid molecule.

6. The method of claim 1, wherein said one or more selectable markers comprises at least one DNA segment encoding an element selected from the group consisting of an antibiotic resistance gene, a gene that encodes a fluorescent protein, an auxotrophic marker, a toxic gene and a phenotypic marker.

7. The method of claim 6, wherein said antibiotic resistance gene is selected from the group consisting of a chloramphenicol resistance gene, an ampicillin resistance gene, a tetracycline resistance gene, a Zeocin resistance gene, a spectinomycin resistance gene and a kanamycin resistance gene.

8. The method of claim 6, wherein said toxic gene is selected from the group consisting of a *ccdB* gene, a gene encoding a *tus* protein, a *kicB* gene, a *sacB* gene, an ASK1 gene, a  $\Phi$ X174 *E* gene and a DpnI gene.

9. The method of claim 1, wherein said first nucleic acid molecule and/or said second nucleic acid molecule further comprises one or more recombination sites.

10. The method of claim 9, wherein said first nucleic acid molecule and/or said second nucleic acid molecule further comprises one or more topoisomerase recognition sites and/or one or more topoisomerases.

11. The method of claim 10, wherein said first nucleic acid molecule and/or said second nucleic acid molecule comprises two or more recombination sites.

12. The method of claim 11, wherein said topoisomerase recognition site, if present, is flanked by said two or more recombination sites.

13. The method of claim 12, wherein said recombination sites are selected from the group consisting of *attB* sites, *attP* sites, *attL* sites, *attR* sites, *lox* sites, *psi* sites, *tnpI* sites, *dif* sites, *cer* sites, *frt* sites, and mutants, variants and derivatives thereof.

14. The method of claim 10, wherein said topoisomerase recognition site, if present, is recognized and bound by a type I topoisomerase.

15. The method of claim 14, wherein said type I topoisomerase is a type IB topoisomerase.

16. The method of claim 15, wherein said type IB topoisomerase is selected from the group consisting of eukaryotic nuclear type I topoisomerase and a poxvirus topoisomerase.

17. The method of claim 1, wherein said expressed interfering RNA is between 35-60 nucleotides in length.

18. The method of claim 17, wherein said expressed interfering RNA forms a hairpin loop.

19. The method of claim 18, wherein said hairpin loop is between 4-8 nucleotides in length.

20. The method of claim 19, wherein said hairpin loop comprises regions of complementarity that are between 18-25 nucleotides in length.

21. A vector comprising:

- (a) one or more toxic genes;
- (b) one or more type IIs restriction enzyme recognition sites; and
- (c) one or more site-specific recombination sites.

22. The vector of claim 21, wherein said type IIs restriction enzyme recognition sites are selected from the group consisting of *Bsa*I, *Bbs*I, *Bbv*II, *Bsm*AI, *Bsp*MI, *Eco*31I, *Bsm*BI, *Bae*I, *Fok*I, *Hga*I, *Sfa*NI and *Sth*132I.

23. The vector of claim 21, wherein said recombination sites are selected from the group consisting of *att*B sites, *att*P sites, *att*L sites, *att*R sites, *lox* sites, *psi* sites, *tnp*I sites, *dif* sites, *cer* sites, *frt* sites, and mutants, variants and derivatives thereof.

24. The vector of claim 21, wherein said vector further comprises one or more topoisomerase recognition sites and/or one or more topoisomerases.

25. The vector of claim 24, wherein said molecule comprises two or more recombination sites.

26. The vector of claim 24, wherein said topoisomerase recognition site, if present, is flanked by said two or more recombination sites.

27. The vector of claim 24, wherein said topoisomerase recognition site, if present, is recognized and bound by a type I topoisomerase.

28. The vector of claim 27, wherein said type I topoisomerase is a type IB topoisomerase.

29. A method of regulating the expression of one or more genes in a transgenic cell or a transgenic animal using interfering RNA, comprising:

(a) identifying one or more target nucleic acid sequences in said cell or animal;

(b) preparing one or more nucleic acid molecules which encode one or more interfering RNAs, wherein said interfering RNAs bind to said one or more target nucleic acid sequences;

(c) combining

(i) one or more first nucleic acid molecules encoding one or more interfering RNAs flanked by one or more first type IIs restriction enzyme recognition sites;

(ii) one or more second nucleic acid molecules comprising one or more selectable markers flanked by one or more second type IIs restriction enzyme recognition sites; and

(iii) one or more site-specific type IIs restriction enzymes; and

(d) incubating said combination under conditions sufficient to join one or more of said one or more nucleic acid molecules encoding one or more interfering RNAs and one or more of said second nucleic acid molecules, thereby producing one or more desired product nucleic acid molecules;

(e) inserting said one or more interfering RNA-containing product nucleic acid molecules into said cell or one or more cells of said animal, under conditions such that said one or more interfering RNAs bind to said one or more target nucleic acid sequences, thereby regulating expression of said one or more genes.

30. The method of claim 29, wherein said expressed interfering RNA is between 35-60 nucleotides in length.

31. The method of claim 30, wherein said expressed interfering RNA forms a hairpin loop.

32. The method of claim 31, wherein said hairpin loop is between 4-8 nucleotides in length.

33. The method of claim 32, wherein said hairpin loop comprises regions of complementarity that are between 18-25 nucleotides in length.

34. The method of claim 29, wherein said regulation results in decreased expression of said one or more genes.

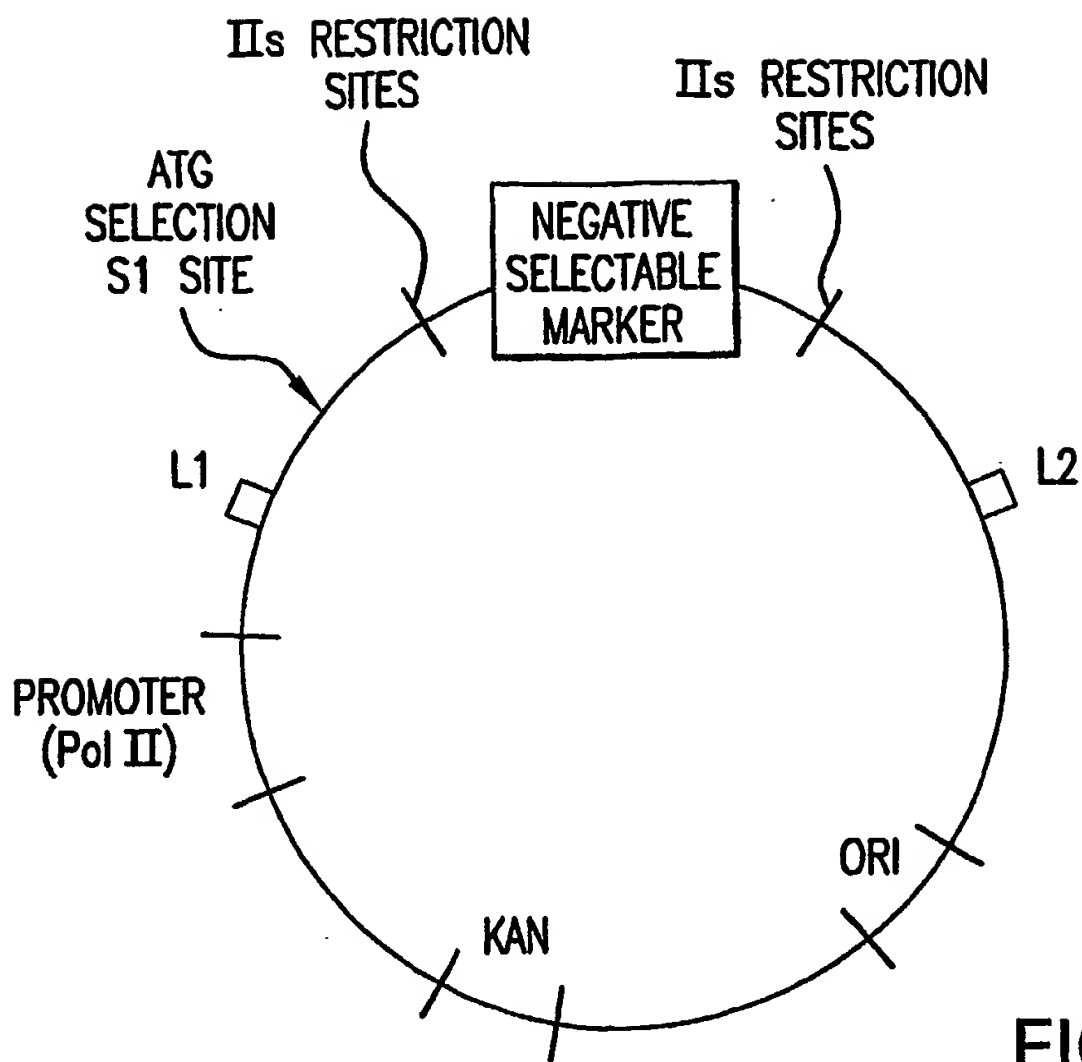


FIG. 1A.

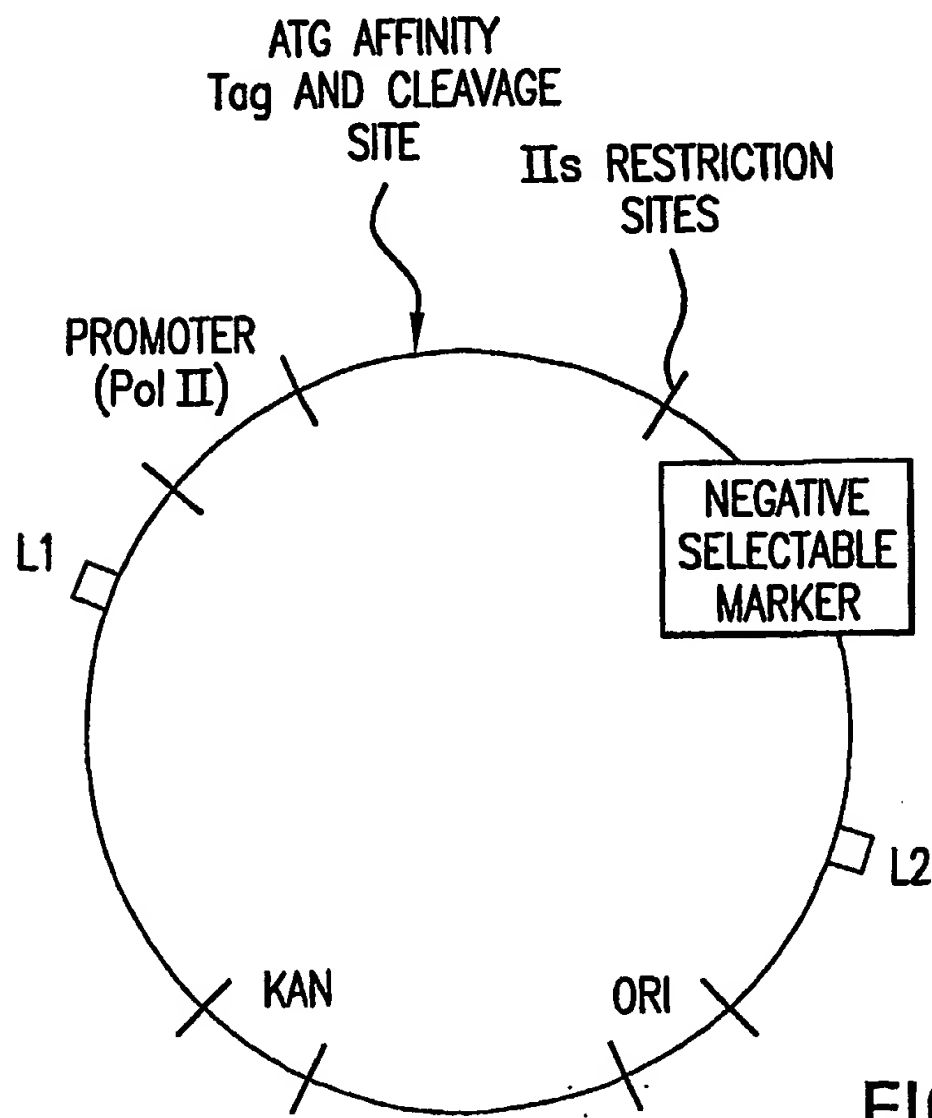


FIG. 1B

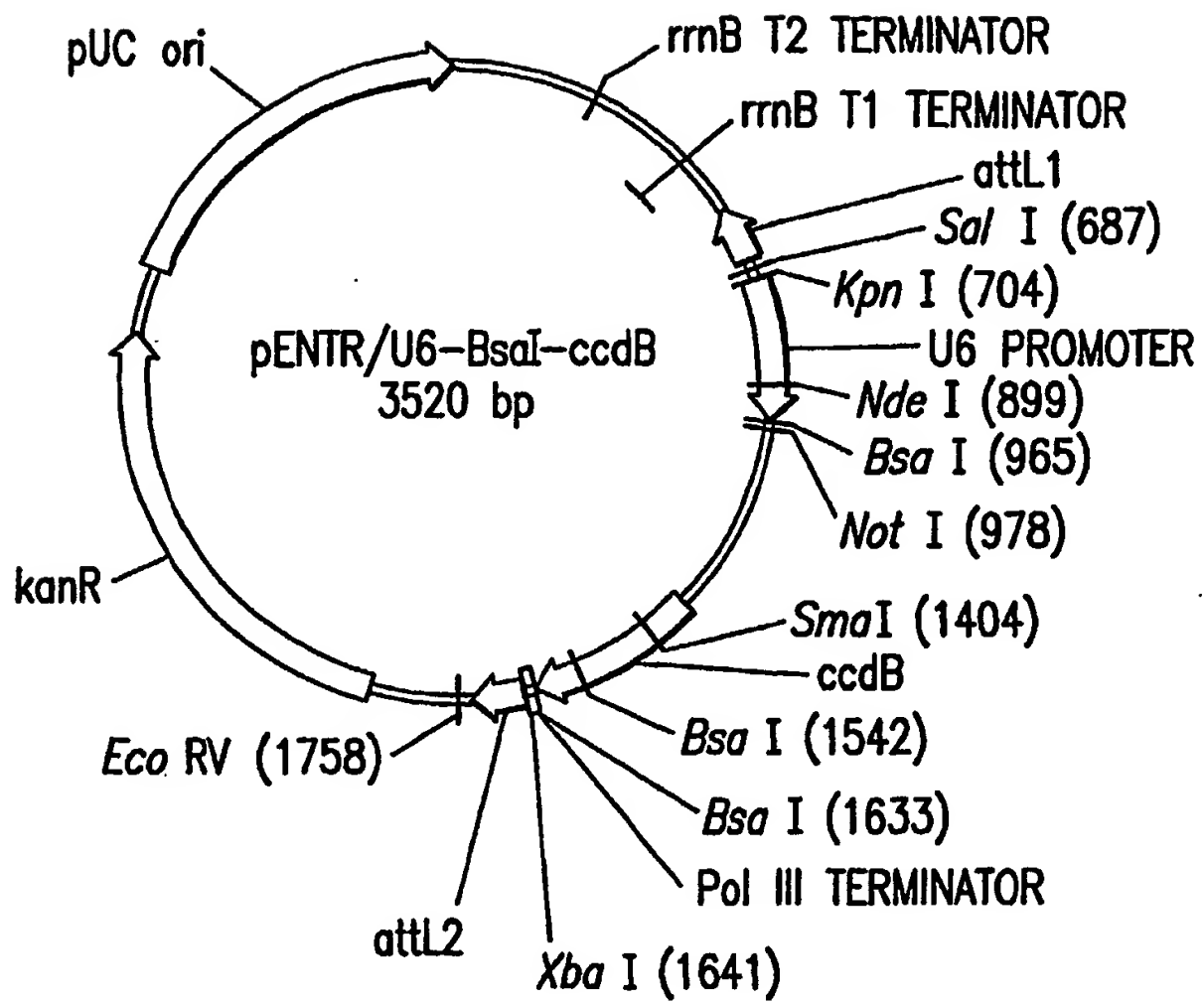
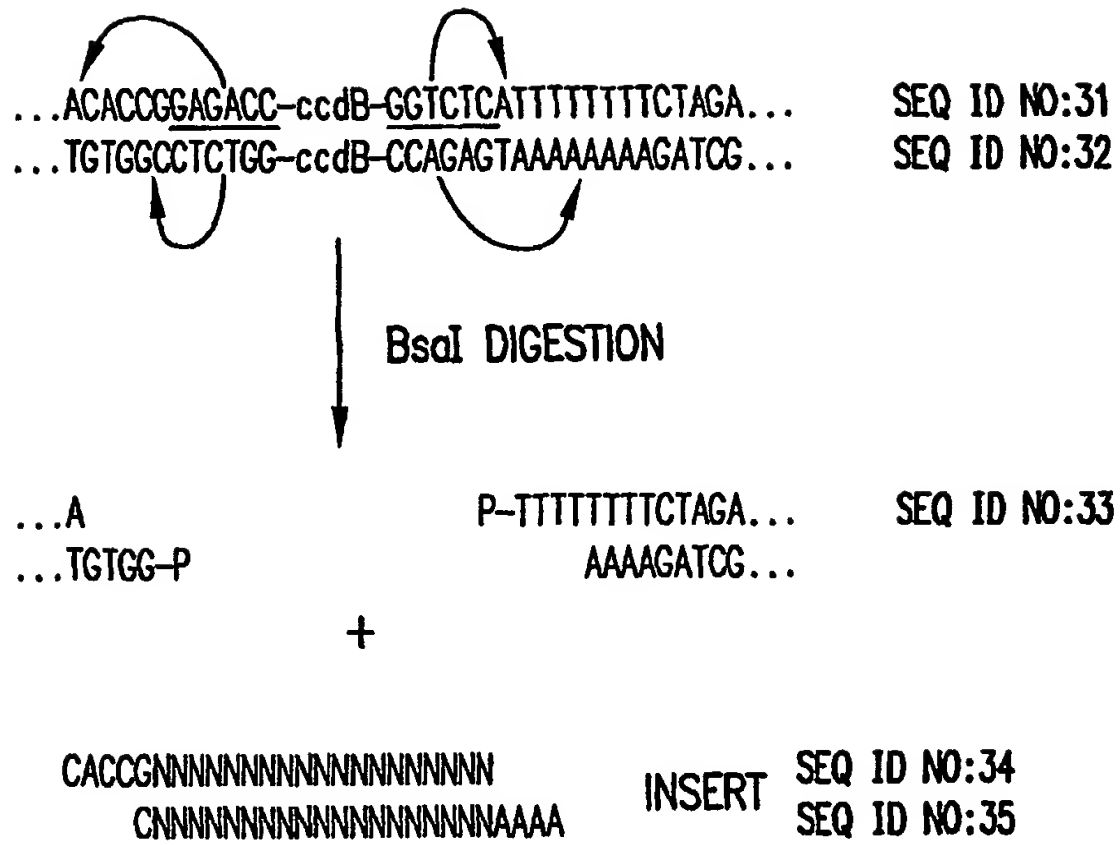
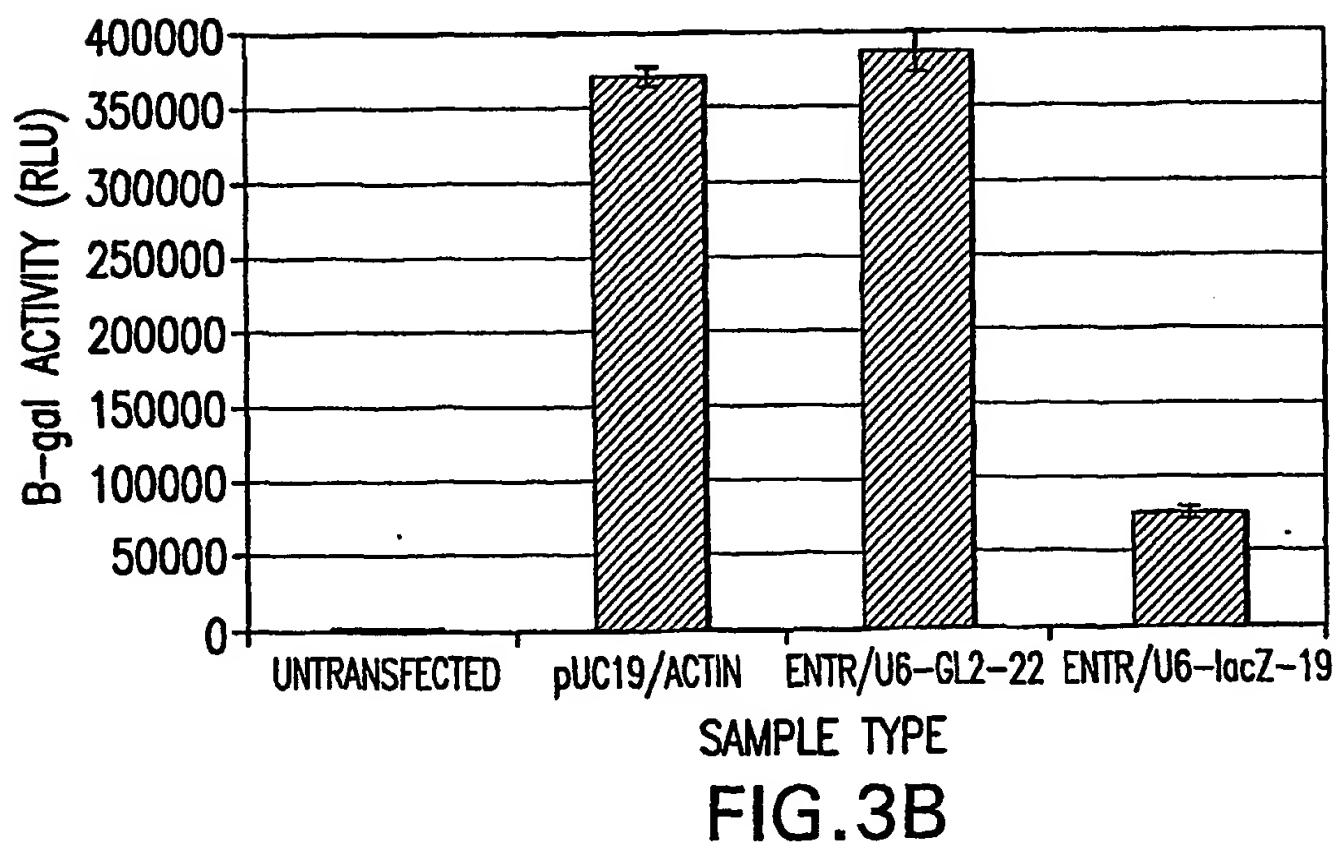
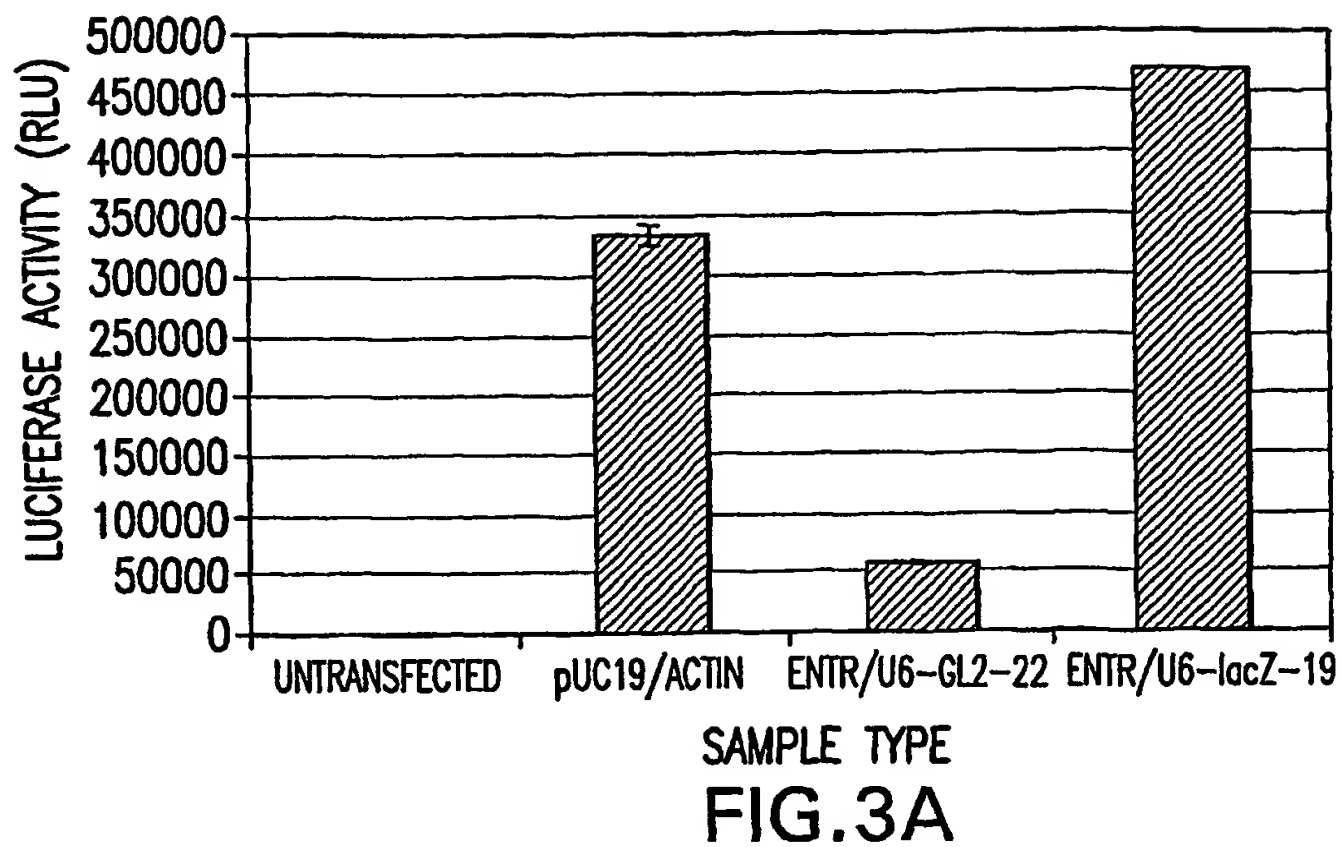


FIG.2A



**FIG. 2B**





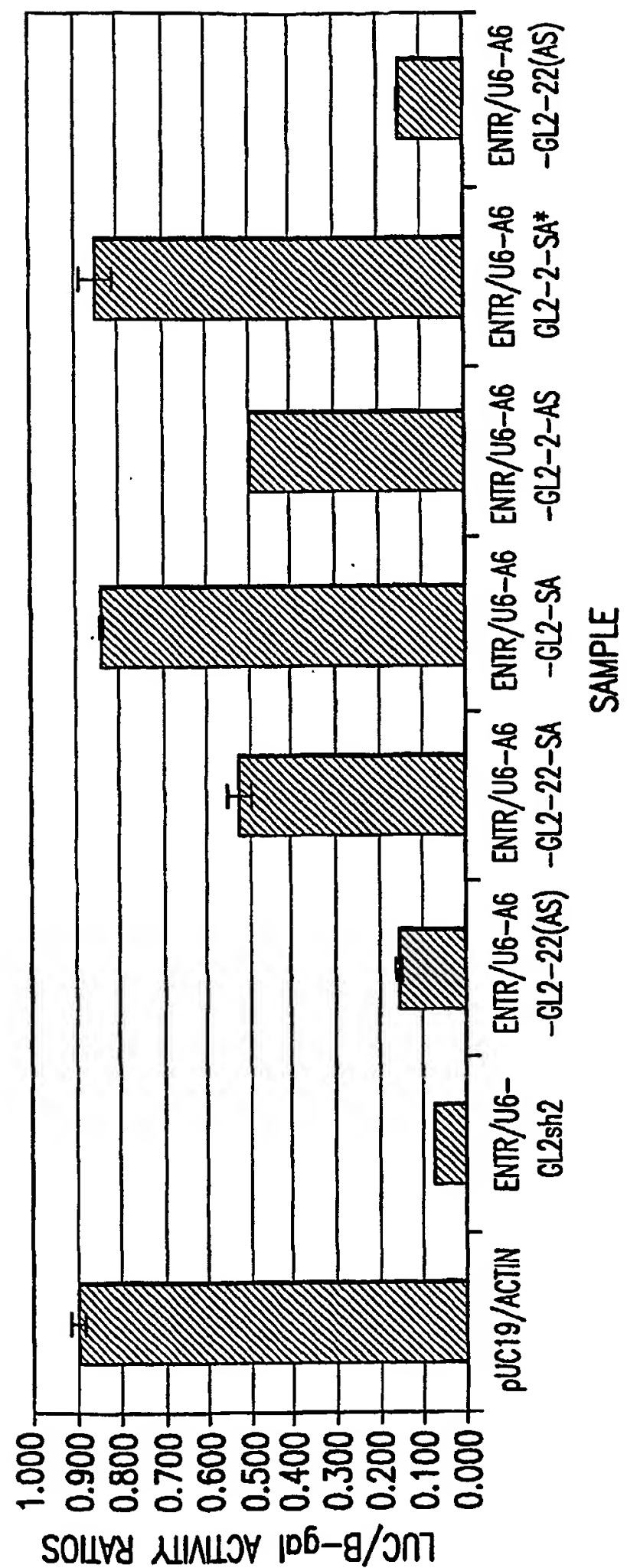
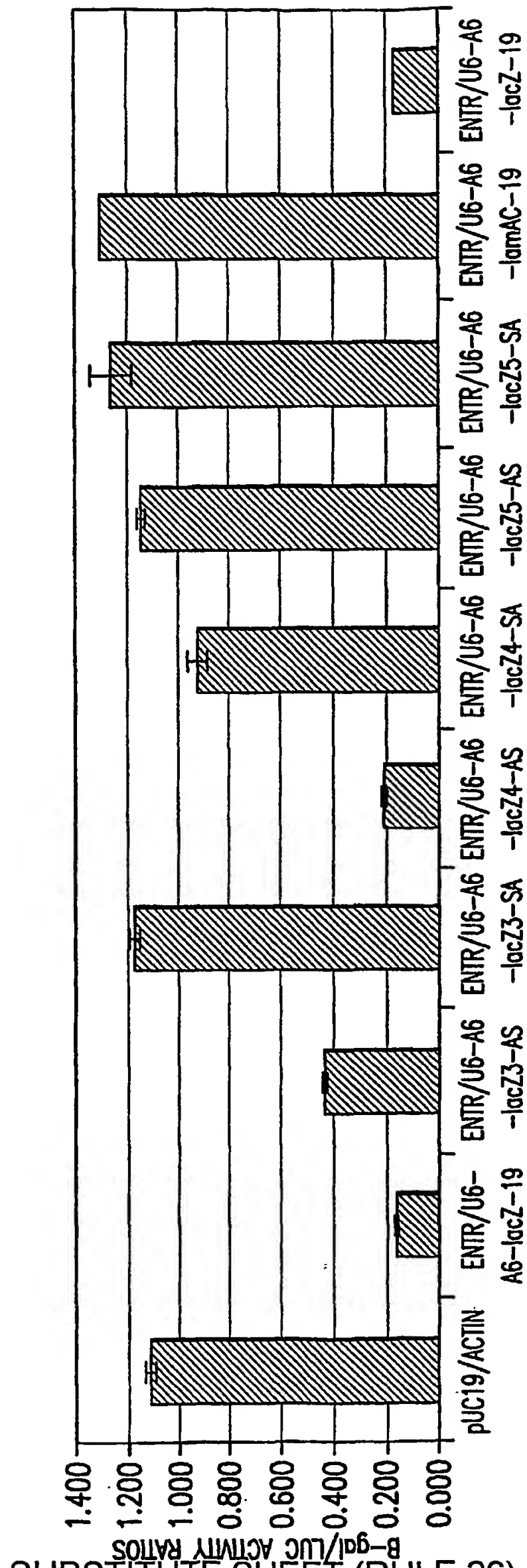


FIG.4A



SAMPLE

FIG.4B

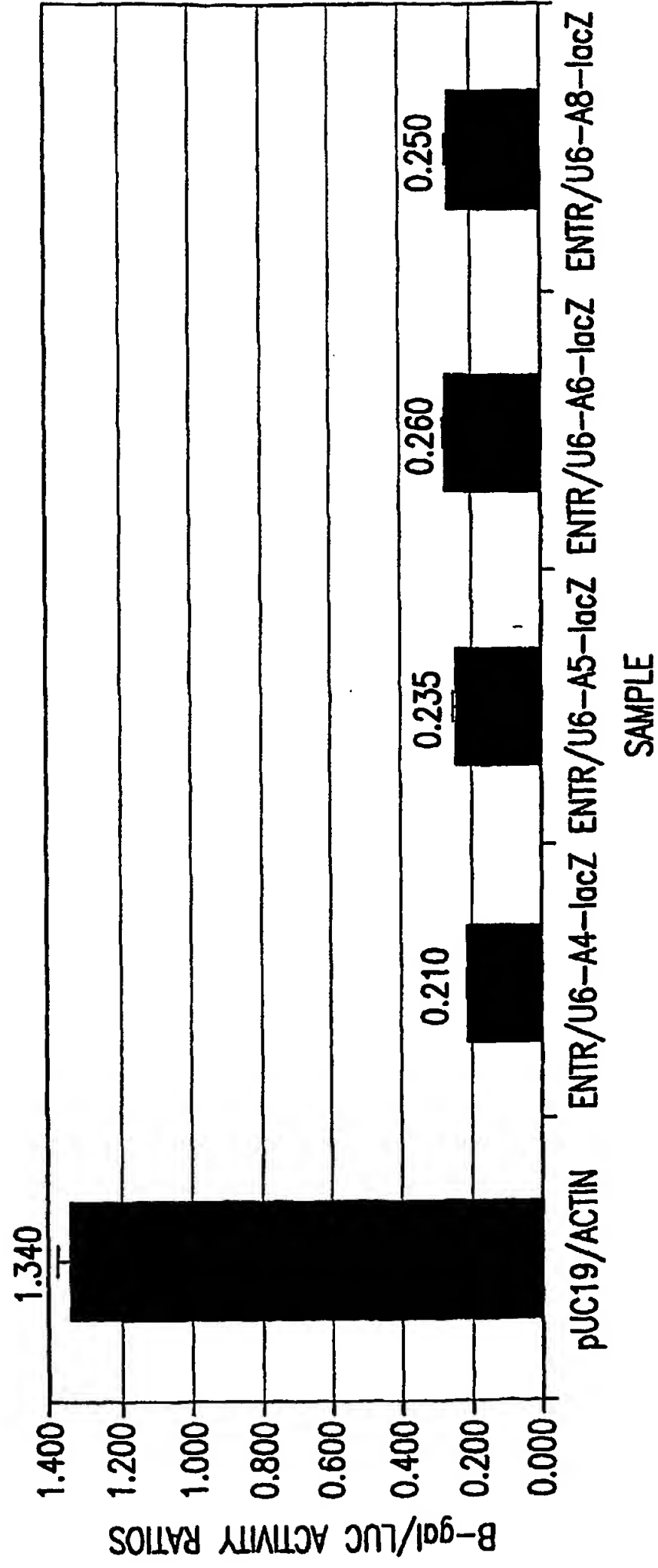


FIG. 5

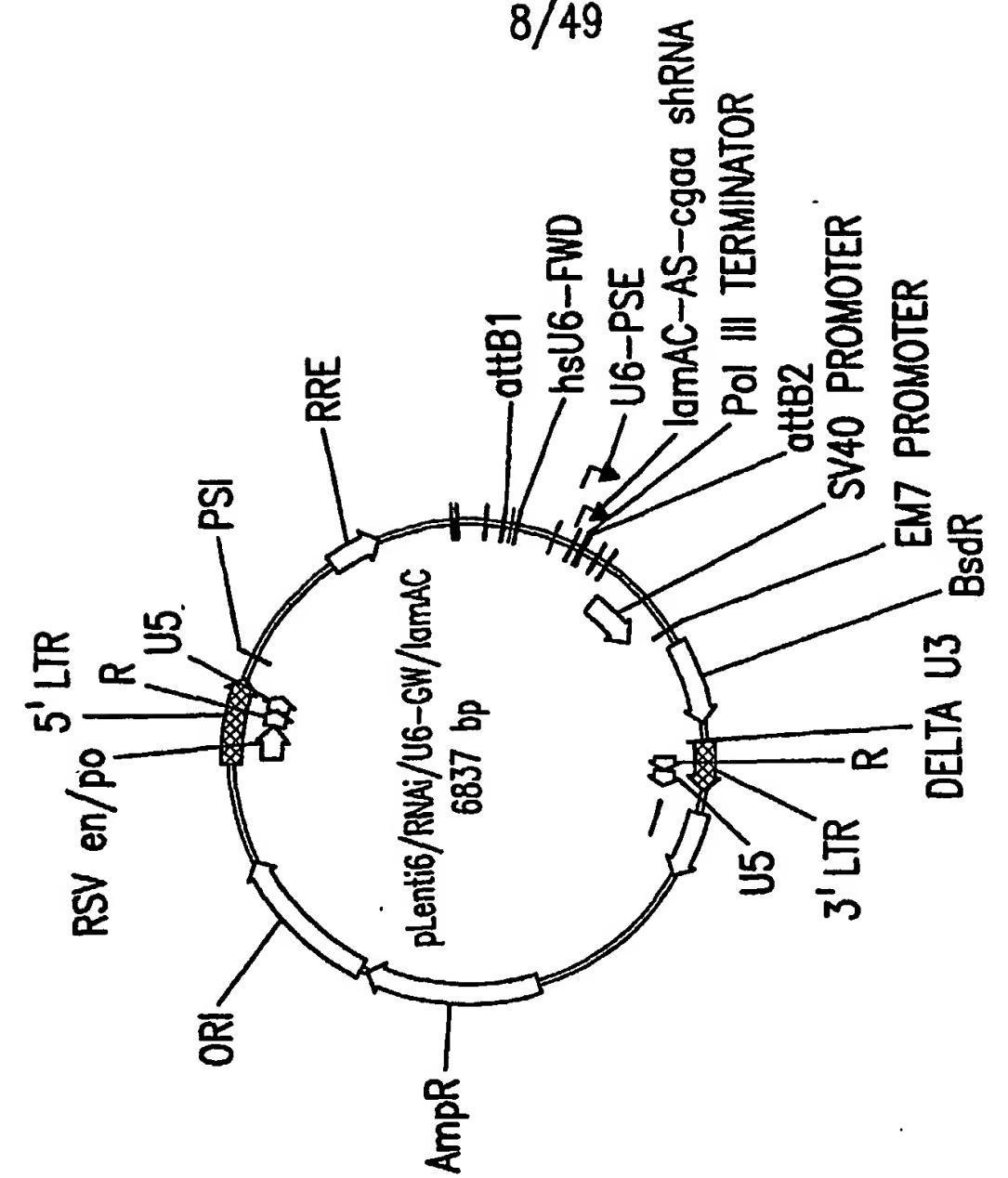


FIG.6B

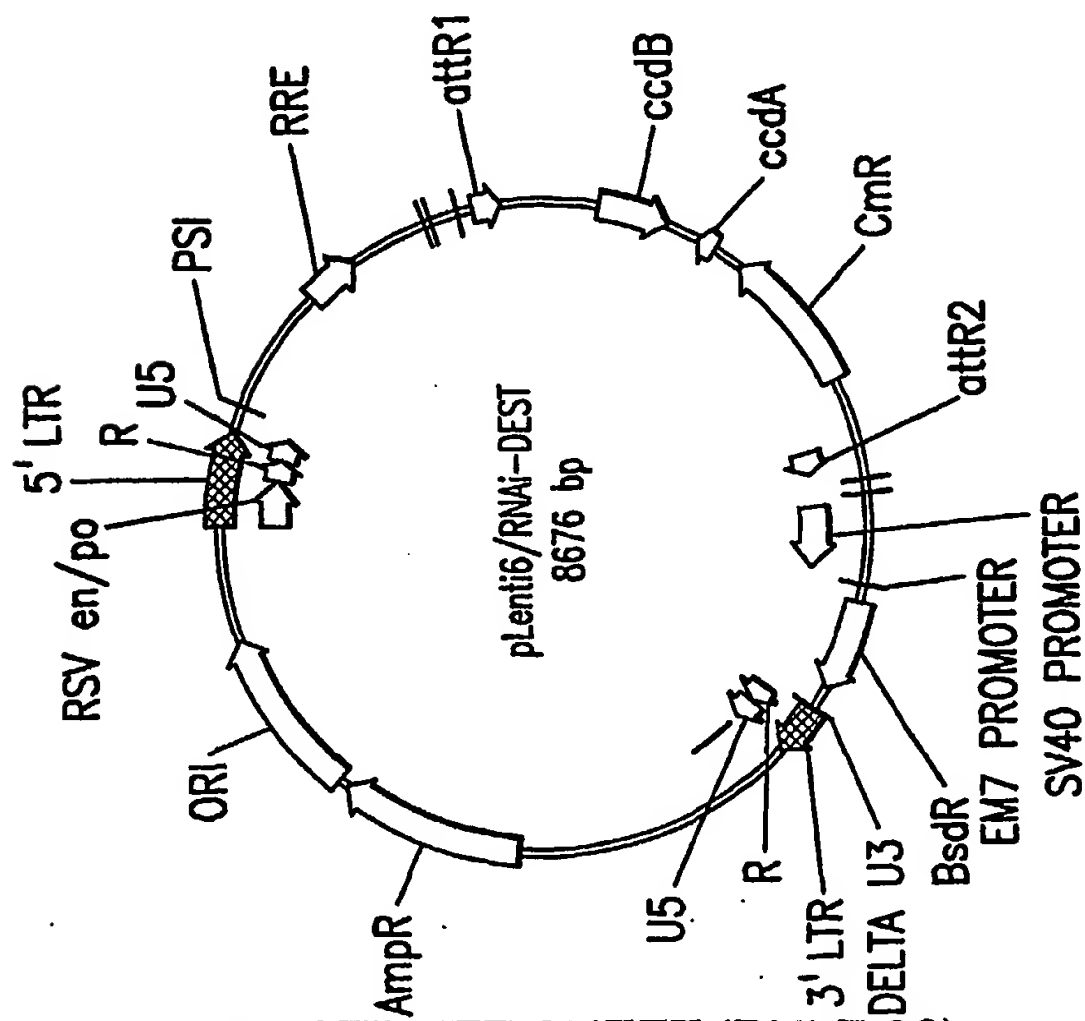


FIG.6A

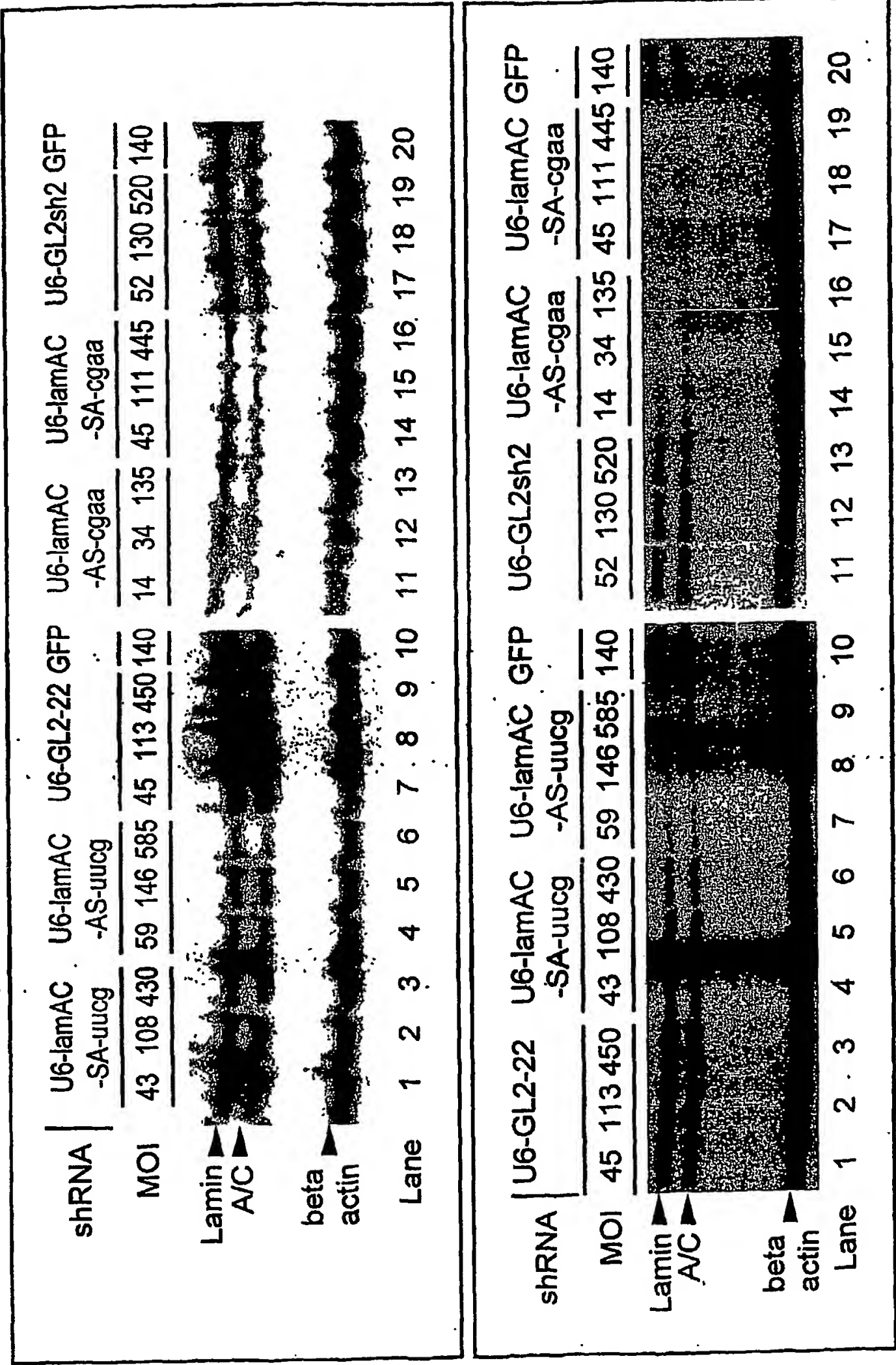


FIG.7

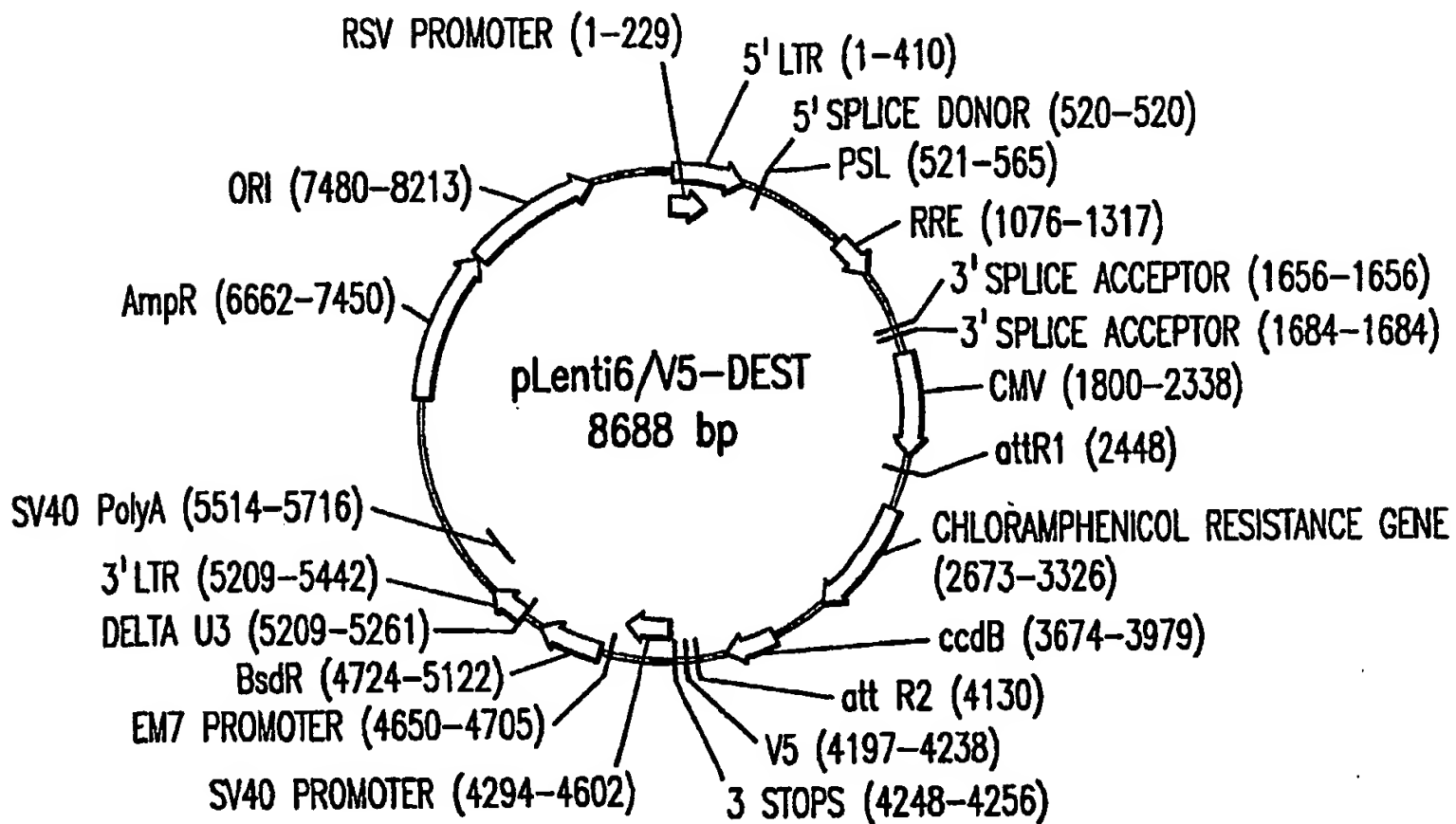


FIG.8A

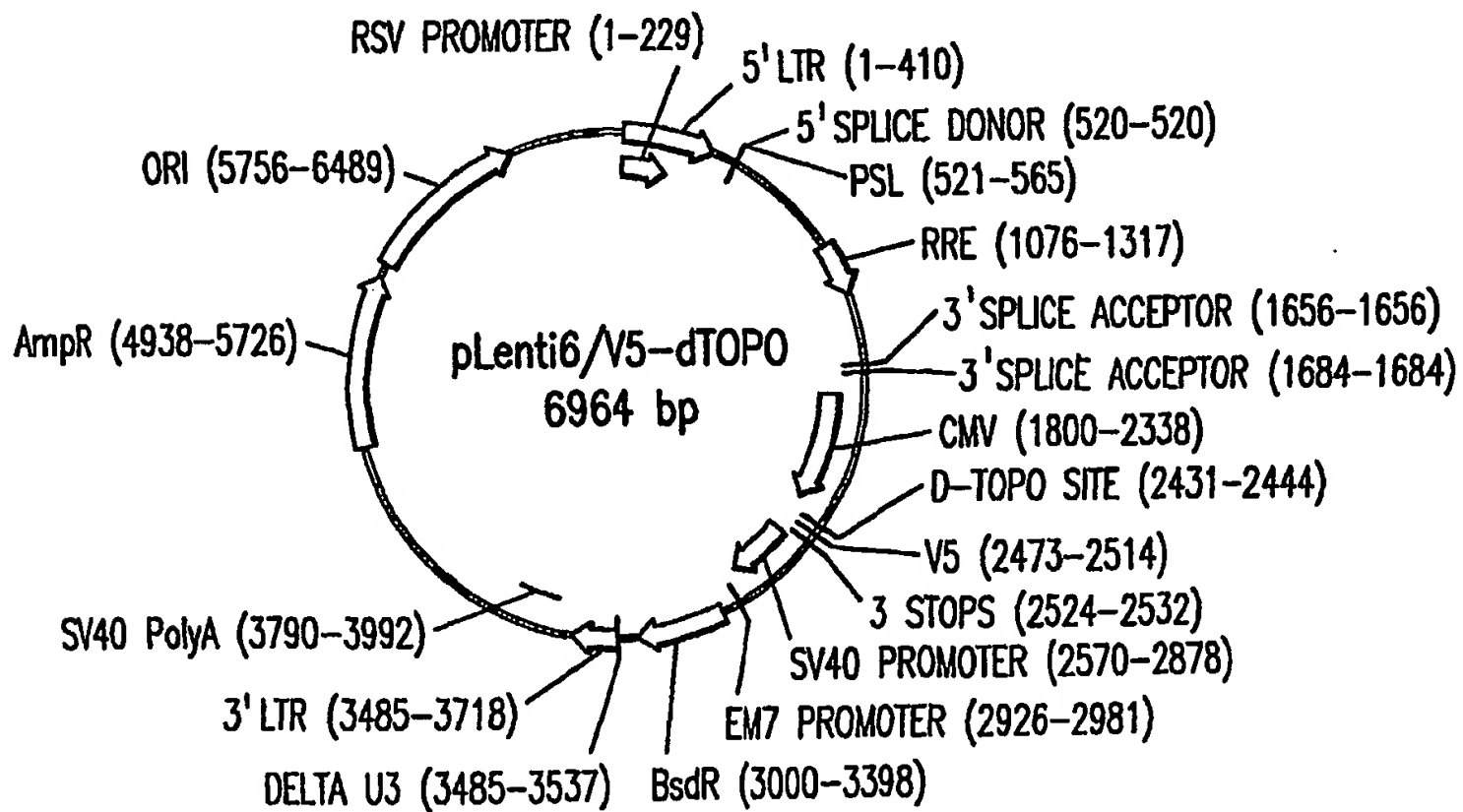


FIG.8B

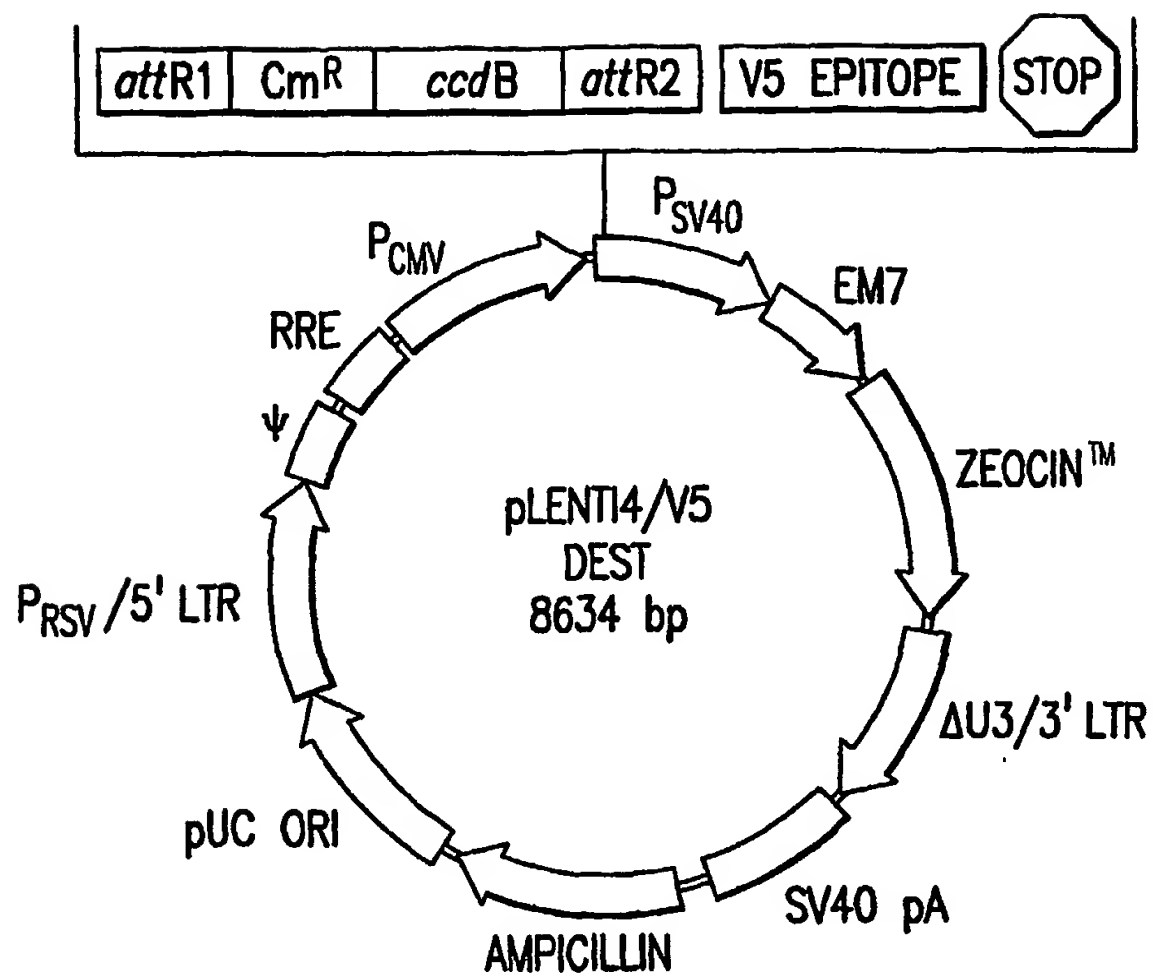


FIG.8C



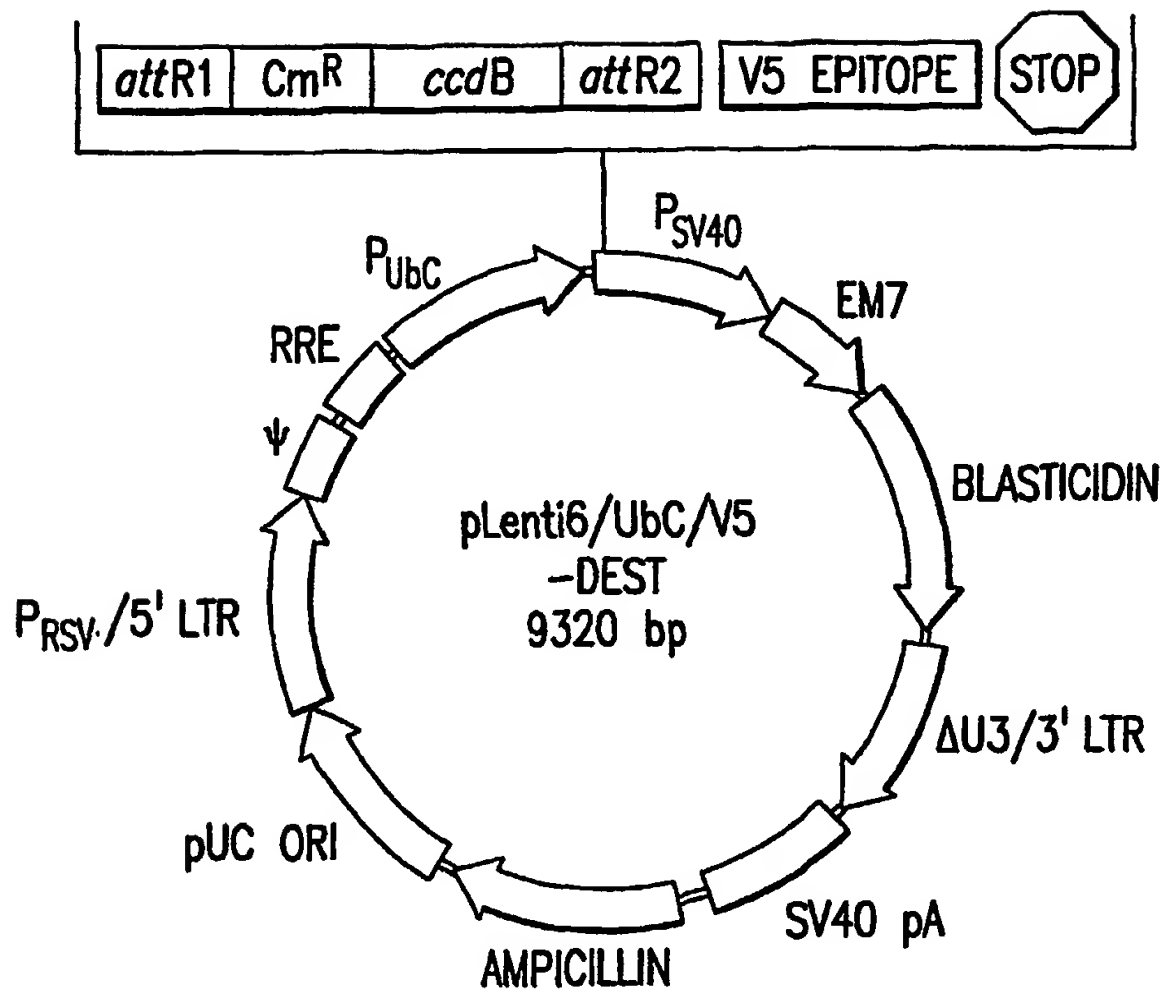


FIG.8D

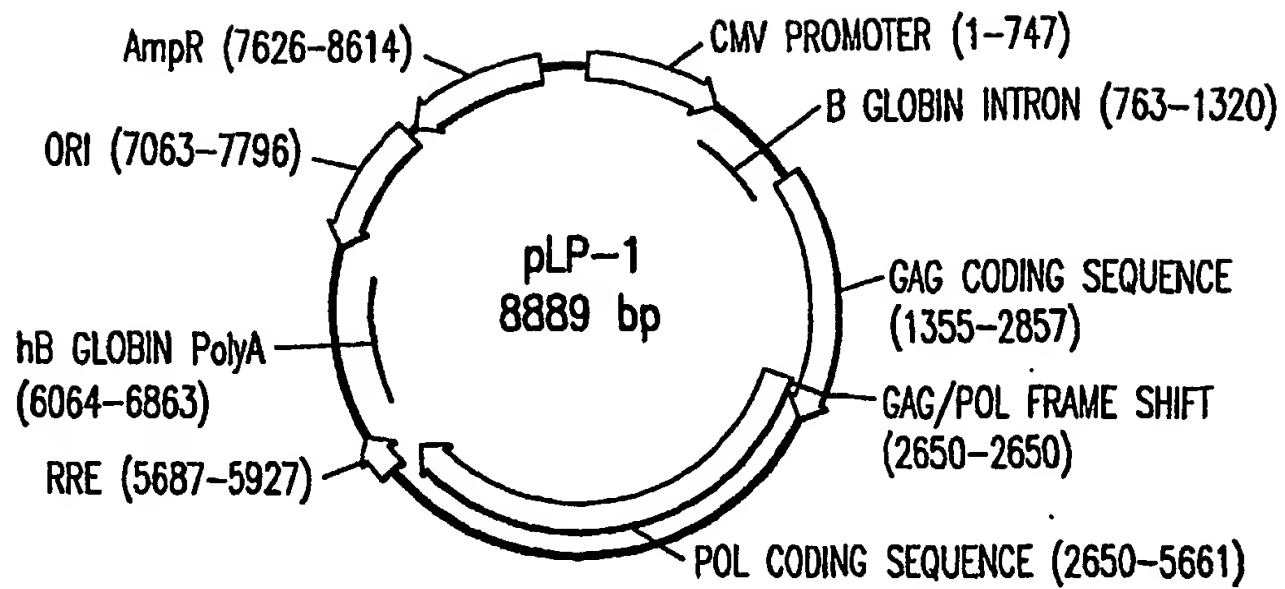


FIG.9A

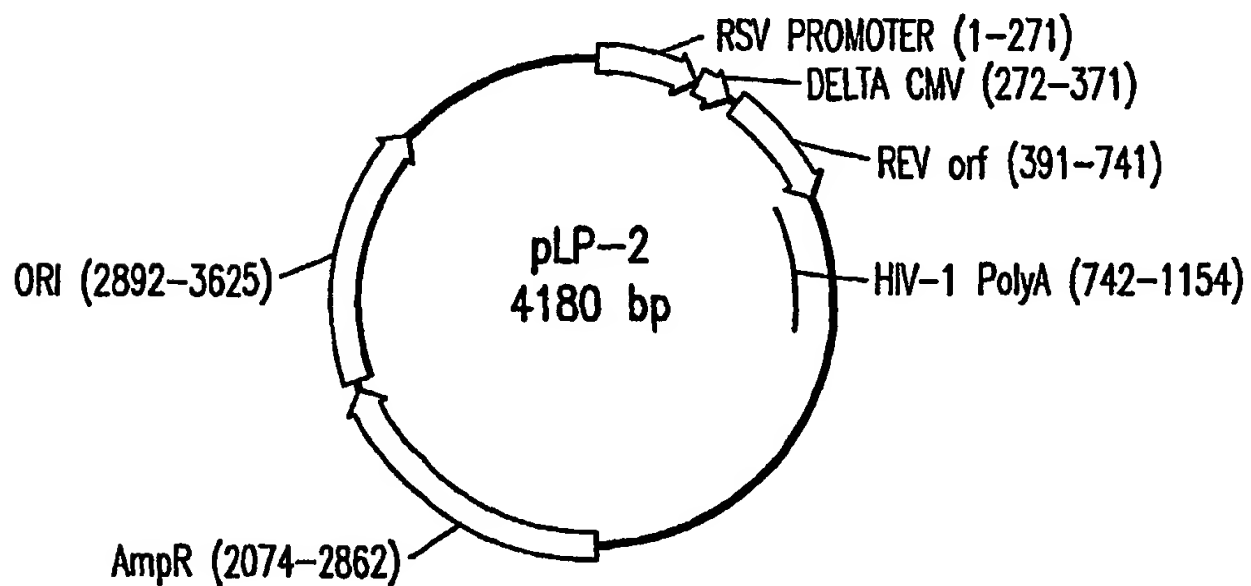


FIG.9B

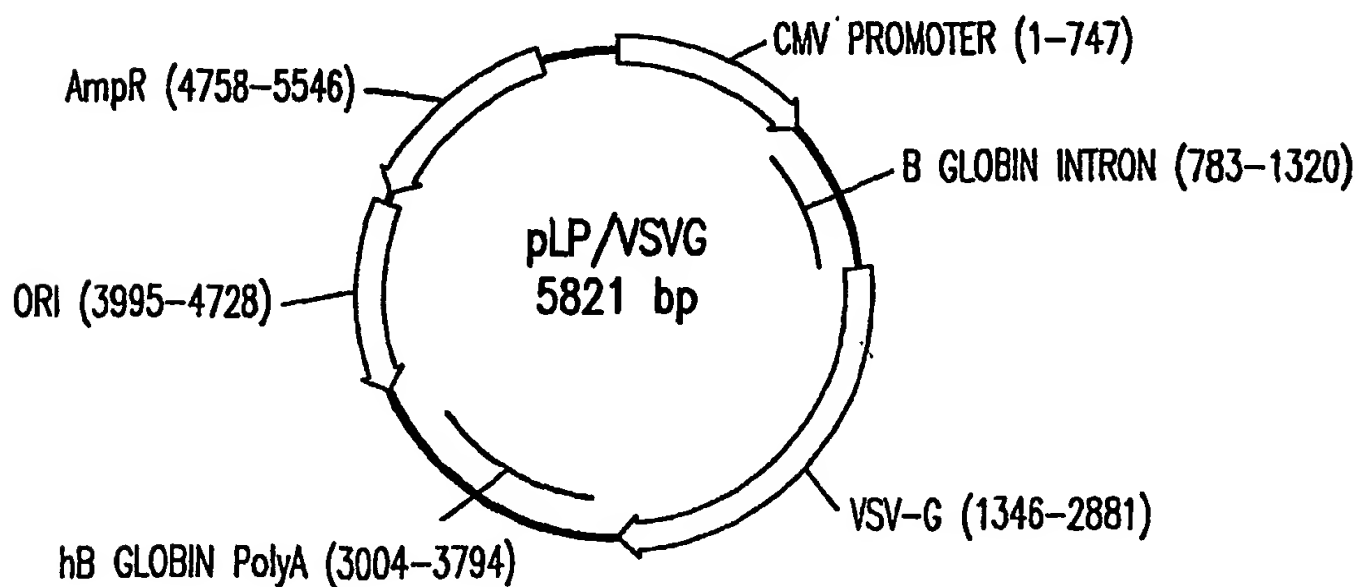


FIG.9C

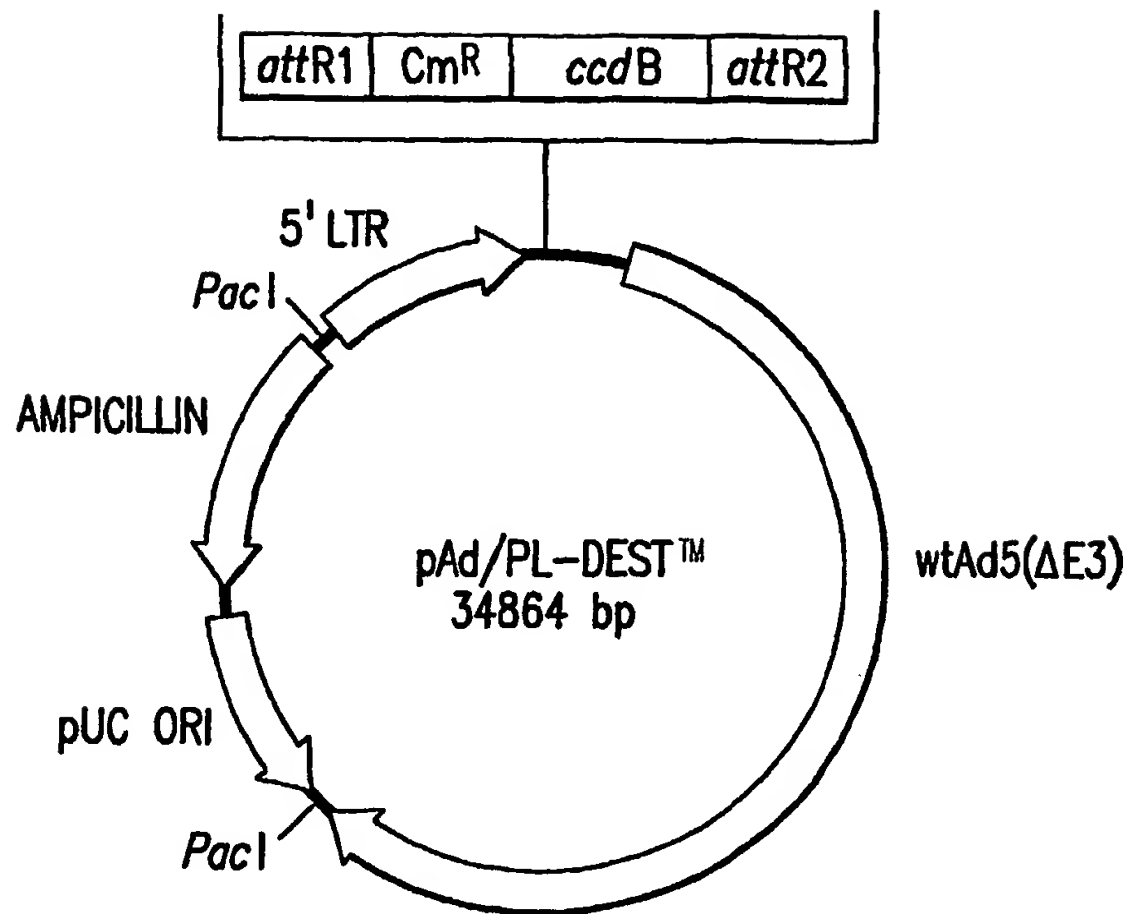


FIG.10

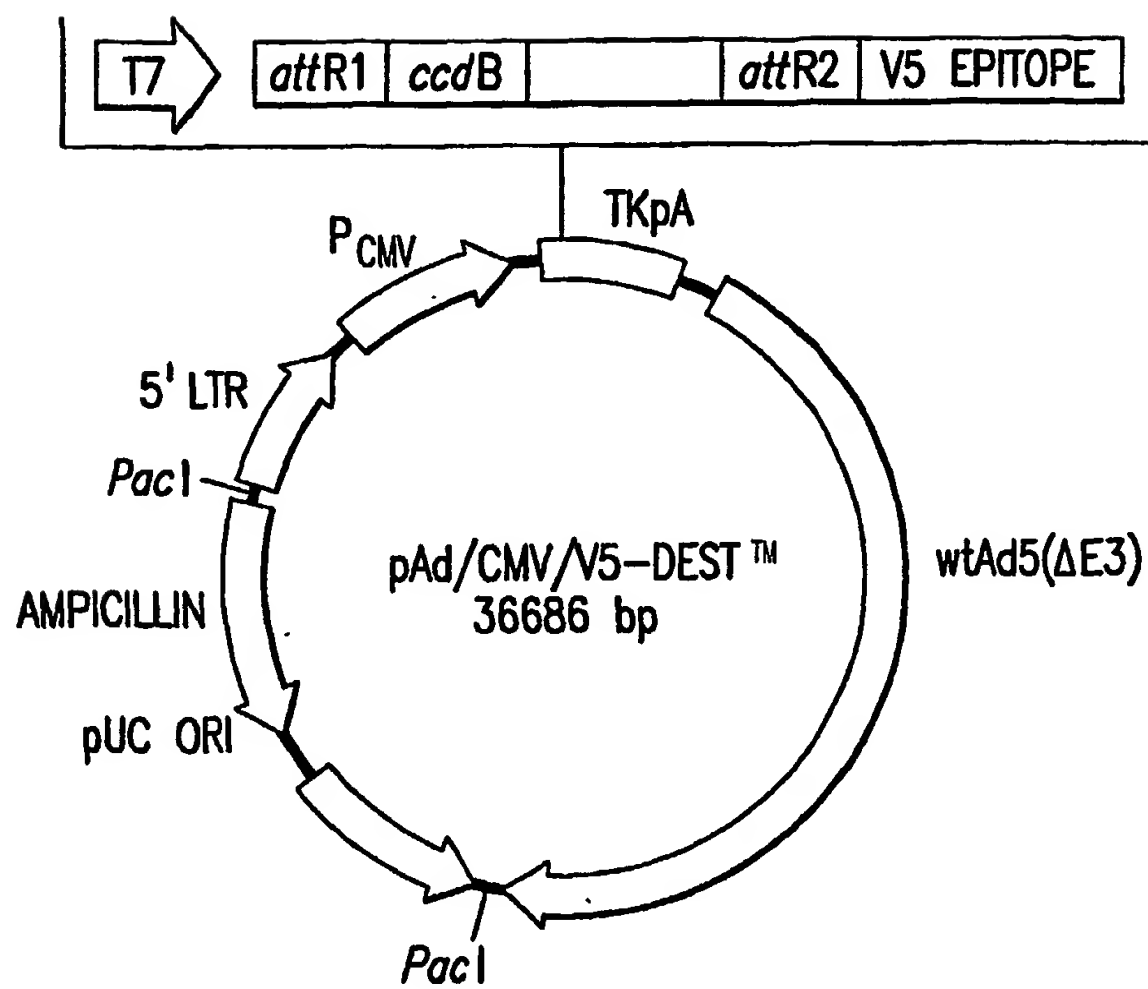


FIG. 11

1 CTTTCCTGCG TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA  
 GAAAGGACGC AATAGGGGAC TAAGACACCT ATTGGCATAA TGGCGGAAAC TCACTCGACT  
  
 61 TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA  
 ATGGCGAGCG GCGTCGGCTT GCTGGCTCGC GTCGCTCAGT CACTCGCTCC TTCGCCTTCT  
  
 121 GCGCCCAATA CGCAAACCGC CTCTCCCCGC GCGTTGGCCG ATTCATTAAT GCAGCTGGCA  
 CGCGGGTTAT GCGTTTGGCG GAGAGGGGCG CGCAACCGGC TAAGTAATTA CGTCGACCGT  
  
 181 CGACAGGTTT CCCGACTGGA AAGCGGGCAG TGAGCGCAAC GCAATTAATA CGCGTACCGC  
 GCTGTCCAAA GGGCTGACCT TTCGCCCCGC ACTCGCGTTG CGTTAATTAT GCGCATGGCG  
  
 241 TAGCCAGGAA GAGTTTGTAG AAACGCAAAA AGGCCATCCG TCAGGATGGC CTTCTGCTTA  
 ATCGGTCCTT CTCAAACATC TTTGCGTTTT TCCGGTAGGC AGTCCTACCG GAAGACGAAT  
 RRN T2 terminator  
 301 GTTTGATGCC TGGCAGTTTA TGGCGGGCGT CCTGCCCGCC ACCCTCCGGG CCGTTGCTTC  
 CAAACTACGG ACCGTCAAAT ACCGCCCCGA GGACGGGCGG TGGGAGGCCG GGCAACGAAG  
  
 361 ACAACGTTCA AATCCGCTCC CGGCGGATTT GTCCTACTCA GGAGAGCGTT CACCGCCAAA  
 TGTTGCAAGT TTAGGCGAGG GCCGCCTAAA CAGGATGAGT CCTCTCGCAA GTGGCTGTTT  
  
 421 CAACAGATAA AACGAAAGGC CCAGTCTTCC GACTGAGCCT TTCGTTTTAT TTGATGCCTG  
 GTTGTCTATT TTGCTTTCG GGTGAGAAGG CTGACTCGGA AAGCAAATA AACTACGGAC  
 RRN T1 terminator  
  
 M13 For(-20)  
 481 GCAGTTCCTT ACTCTCGCGT TAACGCTAGC ATGGATGTTT TCCCAGTCAC GACGTTGTAA  
 CGTCAAGGGA TGAGAGCGCA ATTGCGATCG TACCTACAAA AGGGTCAGTG CTGCAACATT  
M13 For (-20)  
 541 AACGACGGCC AGTCTTAAGC TCGGGCCCCA AATAATGATT TTATTTTGAC TGATAGTGAC  
 TTGCTGCCGG TCAGAATTCG AGCCCGGGGT TTATTACTAA AATAAACTG ACTATCACTG  
 601 CTGTTTCGTTG CAACAAATTG ATGAGCAATG CTTTTTTATA ATGCCAACTT TGTACAAAAA  
 GACAAGCAAC GTTGTTTAAC TACTCGTTAC GAAAAAATAT TACGGTTGAA ACATGTTTTT  
  
 SENSE PRM  
pENTR1a-462F hsU6-1wd  
U6 PROMOTER  
 661 AGCAGGCTTT AAAGGAACCA ATTCAGTCGA CTGGATCCGG TACCAAGGTC GGGCAGGAAG  
 TCGTCCGAAA TTTCTTGGT TAAGTCAGCT GACCTAGGCC ATGGTTCCAG CCCGTCCTTC  
 SENSE PRIV  
hsU6-1w

FIG.12A  
 SUBSTITUTE SHEET (RULE 26)

## U6 PROMOTER

721 AGGGCCTATT TCCCATGATT CCTTCATATT TGCATATACG ATACAAGGCT GTTAGAGAGA  
TCCCGGATAA AGGGTACTAA GGAAGTATAA ACGTATATGC TATGTTCCGA CAATCTCTCT

## U6 PROMOTER

781 TAATTAGAAT TAATTTGACT GTAAACACAA AGATATTAGT ACAAATACG TGACGTAGAA  
ATTAATCTTA ATTAACTGA CATTGTGTGTT TCTATAATCA TGTTTTATGC ACTGCATCTT

## U6 PROMOTER

841 AGTAATAATT TCTTGGGTAG TTTGCAGTTT TAAAATTATG TTTTAAATG GACTATCATA  
TCATTATTAA AGAACCCATC AAACGTCAAA ATTTAATAC AAAATTTTAC CTGATAGTAT

U6-PSE

PROMOTER E

## U6 PROMOTER

901 TGCTTACCGT AACTTGAAAG TATTTGATT TCTTGGCTTT ATATATCTTG TGGAAAGGAC  
ACGAATGGCA TTGAACCTTC ATAAAGCTAA AGAACCGAAA TATATAGAAC ACCTTTCCTG

+1 base transcription starts

U6 PROMOTER

NotI

961 GAAACACCGG AGACCGCGGC CGCTGGATCC GGCTTACTAA AAGCCAGATA ACAGTATGCG  
CTTTGTGGCC TCTGGCGCCG GCGACCTAGG CCGAATGATT TTCGGTCTAT TGTCATACGC

BsaI

1021 TATTTGCGCG CTGATTTTGG CGGTATAAGA ATATATACTG ATATGTATAC CCGAAGTATG  
ATAACGCGC GACTAAAAAC GCCATATTCT TATATATGAC TATACATATG GGCTTCATAC

1081 TCAAAAAGAG GTGTGCTATG AAGCAGCGTA TTACAGTGAC AGTTGACAGC GACAGCTATC  
AGTTTTTCTC CACACGATAC TTCGTCGCAT AATGTCAGTG TCAACTGTCG CTGTCGATAG

1141 AGTTGCTCAA GGCATATATG ATGTCAATAT CTCCGGTCTG GTAAGCACAA CCATGCAGAA  
TCAACGAGTT CCGTATATAC TACAGTTATA GAGGCCAGAC CATTCTGTGT GGTACGTCTT

1201 TGAAGCCCGT CGTCTGCGTG CCGAACGCTG GAAAGCGGAA AATCAGGAAG GGATGGCTGA  
ACTTCGGGCA GCAGACGCAC GGCTTGCGAC CTTTCGCCTT TTAGTCCTTC CCTACCGACT

ccdB

1261 GGTCGCCCCG TTTATTGAAA TGAACGGCTC TTTTGCTGAC GAGAACAGGG ACTGGTGAAA  
CCAGCGGGCC AAATAACTTT ACTTGCCGAG AAAACGACTG CTCTTGTCCT TGACCACTTT

ccdB

1321 TGCAGTTTAA GGTTTACACC TATAAAGAG AGAGCCGTTA TCGTCTGTTT GTGGATGTAC  
ACGTCAAATT CCAAATGTGG ATATTTTCTC TCTCGGCAAT AGCAGACAAA CACCTACATG

ccdB

1381 AGAGTGATAT TATTGACACG CCCGGGCGAC GGATGGTGAT CCCCTGGCC AGTGACAGTC  
TCTCACTATA ATAAGTGTGC GGGCCCGCTG CCTACCACTA GGGGGACCGG TCACGTGCAG

ccdB

1441 TGCTGTCAGA TAAAGTCTCC CGTGAACCTT ACCCGGTGGT GCATATCGGG GATGAAAGCT  
ACGACAGTCT ATTCAGAGG GCACTTGAAA TGGGCCACCA CGTATAGCCC CTACTTTCGA

FIG. 12B

SUBSTITUTE SHEET (RULE 26)

ccdB

---

BsaI

1501 GGC GCATGAT GACCACCGAT ATGGCCAGTG TGCCGGTCTC CGTTATCGGG GAAGAAGTGG  
CCGCGTACTA CTGGTGGCTA TACCGGTCAC ACGGCCAGAG GCAATAGCCC CTTCTTCACC

---

ccdB

1561 CTGATCTCAG CCACCGCGAA AATGACATCA AAAACGCCAT TAACCTGATG TTCTGGGGAA  
GACTAGAGTC GGTGGCGCTT TTAGTGTAGT TTTTGCGGTA ATTGGACTAC AAGACCCCTT

ccdB            Pol III terminator

BsaI

1621 TATAAGGTCT CATTTTTTTT CTAGACCCAG CTTTCTTGTA CAAAGTTGGC ATTATAAGAA  
ATATTCCAGA GTAAAAAAA GATCTGGGTC GAAAGAACAT GTTCAACCG TAATATTCTT

1681 AGCATTGCTT ATCAATTTGT TGCAACGAAC AGGTCACTAT CAGTCAAAAT AAAATCATT  
TCGTAACGAA TAGTTAAACA ACGTTGCTTG TCCAGTGATA GTCAGTTTAA TTTTAGTAAT

M13 Rev

1741 TTTGCCATCC AGCTGATATC CCCTATAGTG AGTCGTATTA CATGGTCATA GCTGTTTCCT  
AAACGGTAGG TCGACTATAG GGGATATCAC TCAGCATAAT GTACCAGTAT CGACAAAGGA

M13 Rev

1801 GGCAGCTCTG GCCCGTGTCT CAAAATCTCT GATGTTACAT TGCACAAGAT AAAAATATAT  
CCGTCGAGAC CGGGCACAGA GTTTTAGAGA CTACAATGTA ACGTGTTCTA TTTTATATA

kanR

1861 CATCATGAAC AATAAACTG TCTGCTTACA TAAACAGTAA TACAAGGGGT GTTATGAGCC  
GTAGTACTTG TTATTTTGAC AGACGAATGT ATTTGTCATT ATGTTCCCA CAATACTCGG

kanR

1921 ATATTCAACG GGAAACGTCG AGGCCGCGAT TAAATTCCAA CATGGATGCT GATTTATATG  
TATAAGTTGC CCTTTGCAGC TCCGGCGCTA ATTTAAGGTT GTACCTACGA CTAAATATAC

kanR

1981 GGTATAAATG GGCTCGCGAT AATGTCGGGC AATCAGGTGC GACAATCTAT CGCTTGTATG  
CCATATTTAC CCGAGCGCTA TTACAGCCCG TTAGTCCACG CTGTTAGATA GCGAACATAC

kanR

2041 GGAAGCCCGA TCGCCAGAG TTGTTTCTGA AACATGGCAA AGGTAGCGTT GCCAATGATG  
CCTTCGGGCT ACGCGGTCTC AACAAAGACT TTGTACCGTT TCCATCGCAA CGGTTACTAC

kanR

2101 TTACAGATGA GATGGTCAGA CTAACTGGC TGACGGAATT TATGCCTCTT CCGACCATCA  
AATGTCTACT CTACCAGTCT GATTTGACCG ACTGCCTTAA ATACGGAGAA GGCTGGTAGT

kanR

2161 AGCATTTTAT CCGTACTCCT GATGATGCAT GGTTACTCAC CACTGCGATC CCCGAAAAA  
TCGTAAAATA GGCATGAGGA CTACTACGTA CCAATGAGTG GTGACGCTAG CGGCCTTTTT

FIG.12C

SUBSTITUTE SHEET (RULE 26)

kanR

2221 CAGCATTCCA GGTATTAGAA GAATATCCTG ATTCAGGTGA AAATATTGTT GATGCGCTGG  
GTCGTAAGGT CCATAATCTT CTTATAGGAC TAAGTCCACT TTTATAACAA CTACGCGACC

kanR

2281 CAGTGTTCTT GCGCCGGTTG CATTGATTG CTGTTTGTA TGTCTCTTT AACAGCGATC  
GTCACAAGGA CGCGGCCAAC GTAAGCTAAG GACAAACATT AACAGGAAAA TTGTCGCTAG

kanR

2341 GCGTATTTTCG TCTCGCTCAG GCGCAATCAC GAATGAATAA CGGTTTGGTT GATGCGAGTG  
CGCATAAAGC AGAGCGAGTC CGCGTTAGTG CTTACTTATT GCCAAACCAA CTACGCTCAC

kanR

2401 ATTTTGATGA CGAGCGTAAT GGCTGGCCTG TTGAACAAGT CTGGAAAGAA ATGCATAAAC  
TAAACTACT GCTCGCATT CCGACCGGAC AACTTGTTCA GACCTTTCTT TACGTATTTG

kanR

2461 TTTTGCCATT CTCACCGGAT TCAGTCGTCA CTCATGGTGA TTTCTCACTT GATAACCTTA  
AAAACGGTAA GAGTGGCCTA AGTCAGCAGT GAGTACCACT AAAGAGTGAA CTATTGGAAT

kanR

2521 TTTTGACGA GGGGAAATTA ATAGGTGTGA TTGATGTTGG ACGAGTCGGA ATCGCAGACC  
AAAACTGCT CCCCTTTAAT TATCCAACAT AACTACAACC TGCTCAGCCT TAGCGTCTGG

kanR

2581 GATACCAGGA TCTTGCCATC CTATGGAAC GCCTCGGTGA GTTTTCTCCT TCATTACAGA  
CTATGGTCCT AGAACGGTAG GATACCTTGA CGGAGCCACT CAAAAGAGGA AGTAATGTCT

kanR

2641 AACGGCTTTT TCAAAAATAT GGTATTGATA ATCCTGATAT GAATAAATTG CAGTTTCATT  
TTGCCGAAAA AGTTTTTATA CCATAACTAT TAGGACTATA CTTATTTAAC GTCAAAGTAA

kanR

2701 TGATGCTCGA TGAGTTTTTC TAATCAGAAT TGGTTAATTG GTTGTAACAC TGGCAGAGCA  
ACTACGAGCT ACTCAAAAAG ATTAGTCTTA ACCAATTAAC CAACATTGTG ACCGTCTCGT

2761 TTACGCTGAC TTGACGGGAC GGCGCAAGCT CATGACCAA ATCCCTTAAC GTGAGTTACG  
AATGCGACTG AACTGCCCTG CCGCGTTTCA GTACTGGTTT TAGGGAATTG CACTCAATGC

pUC ori

2821 CGTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT  
GCAGCAAGGT GACTCGCAGT CTGGGGCATC TTTTCTAGTT TCCTAGAAGA ACTCTAGGAA

pUC ori

2881 TTTTCTGCG CGTAATCTGC TGCTTGCAA CAAAAAACC ACCGCTACCA GCGGTGGTTT  
AAAAAGACGC GCATTAGACG ACGAACGTTT GTTTTTTTGG TGGCGATGGT CGCCACCAA

pUC ori

2941 GTTTGCCGGA TCAAGAGCTA CCAACTCTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC  
CAAACGGCCT AGTTCTCGAT GGTTGAGAAA AAGGCTTCCA TTGACCGAAG TCGTCTCGCG



pUC ori

3001 AGATACCAA TACTGTCCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG  
TCTATGGTTT ATGACAGGAA GATCACATCG GCATCAATCC GGTGGTGAAG TTCTTGAGAC

pUC ori

3061 TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG  
ATCGTGGCGG ATGTATGGAG CGAGACGATT AGGACAATGG TCACCGACGA CGGTCACCGC

pUC ori

3121 ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT  
TATTCAGCAC AGAATGGCCC AACCTGAGTT CTGCTATCAA TGGCCTATTG CGCGTCGCCA

pUC ori

3181 CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGGA GCGAACGACC TACACCGAAC  
GCCCCACTTG CCCCCAAGC ACGTGTGTCG GGTCGAACCT CGCTTGCTGG ATGTGGCTTG

pUC ori

3241 TGAGATACCT ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCGAAGGG AGAAAGGCGG  
ACTCTATGGA TGTCGCACTC GTAACCTTTT CGCGGTGCGA AGGGCTTCCC TCTTCCGCC

pUC ori

3301 ACAGGTATCC GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG  
TGTCCATAGG CCATTCGCCG TCCAGCCTT GTCCTCTCGC GTGCTCCCTC GAAGGTCCCC

pUC ori

3361 GAAACGCCTG GTATCTTTAT AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT  
CTTTGCGGAC CATAGAAATA TCAGGACAGC CAAAGCGGT GGAGACTGAA CTCGCAGCTA

pUC ori

3421 TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCTTTT  
AAAACACTAC GAGCAGTCCC CCGCCTCGG ATACCTTTTT GCGGTCGTTG CGCCGGAAAA

pUC ori

3481 TACGGTTCCT GGCCTTTTGC TGGCCTTTTG CTCACATGTT  
ATGCCAAGGA CCGGAAAACG ACCGGAAAAC GAGTGTACAA

FIG. 12E

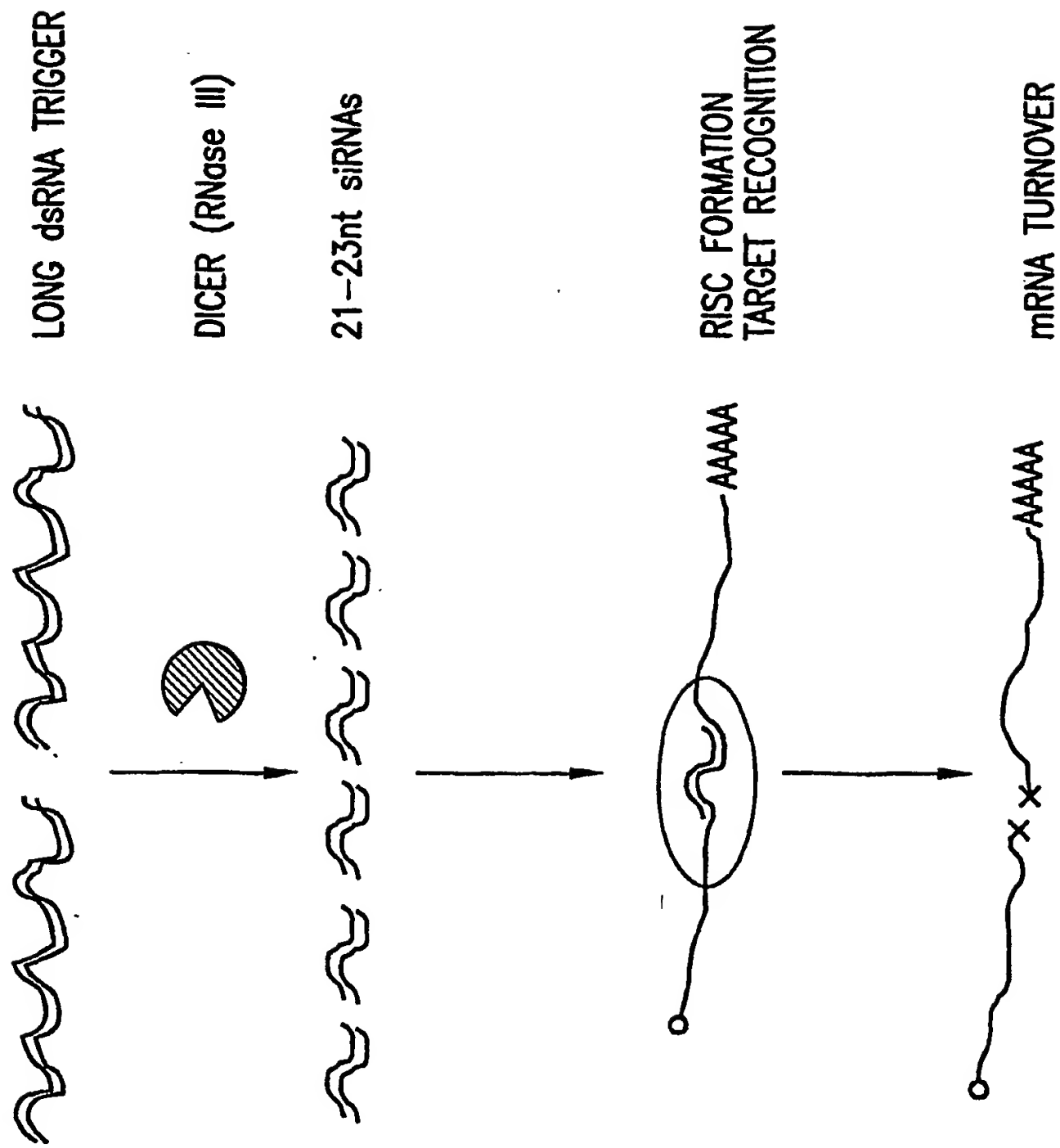


FIG.13

siRNA/RISC = miRNA/miRNP ?

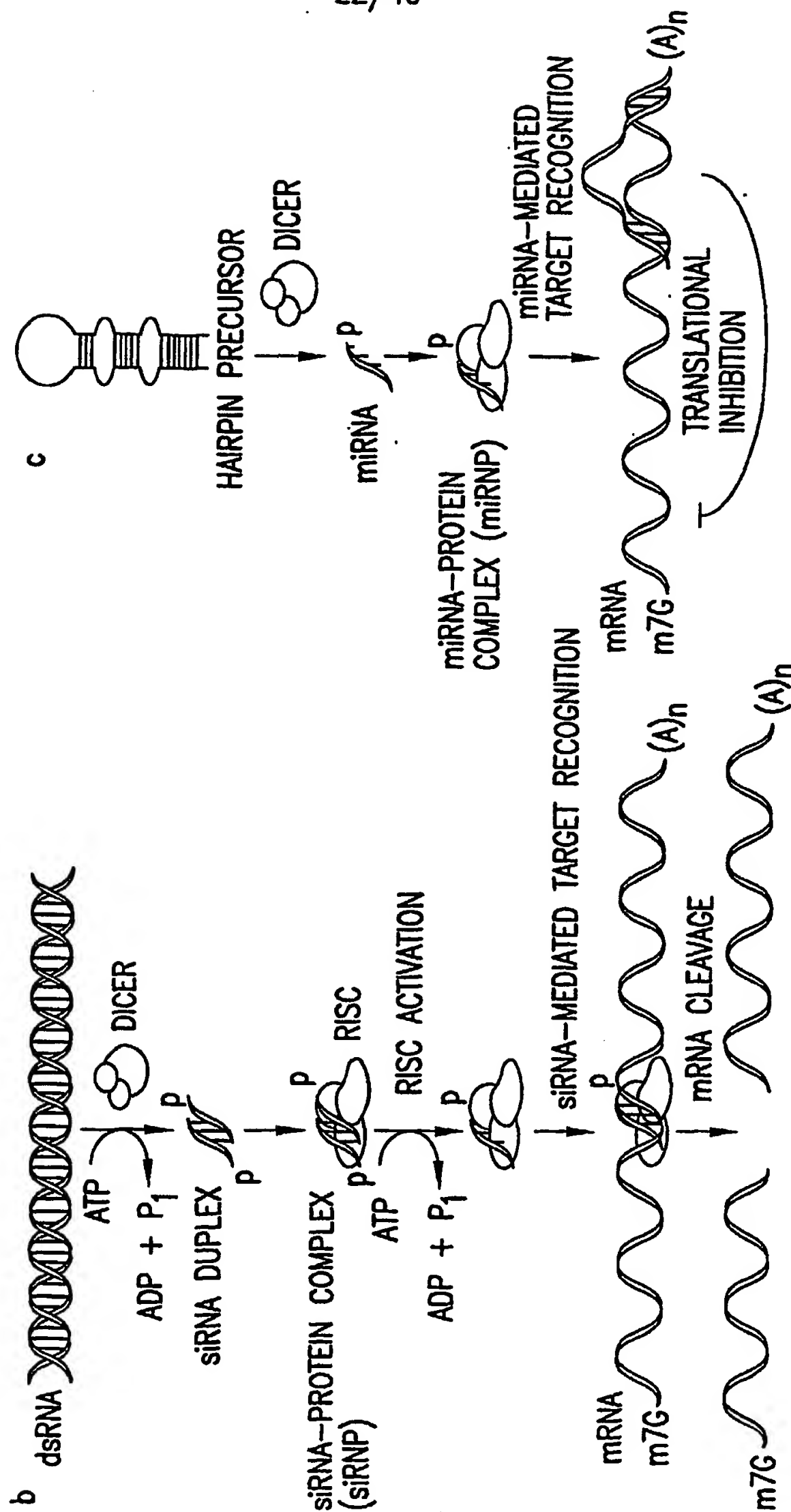


FIG.14

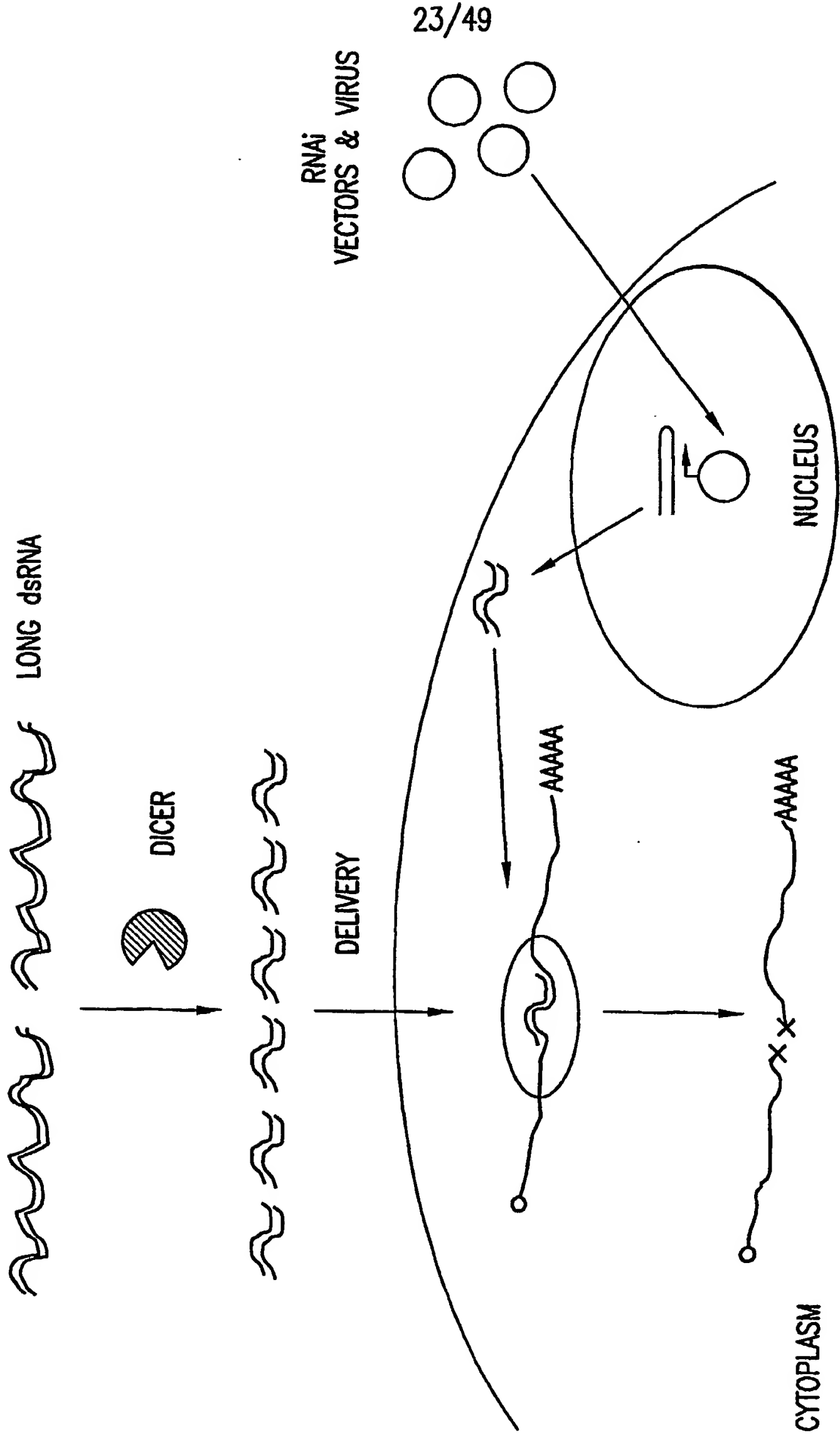


FIG.15

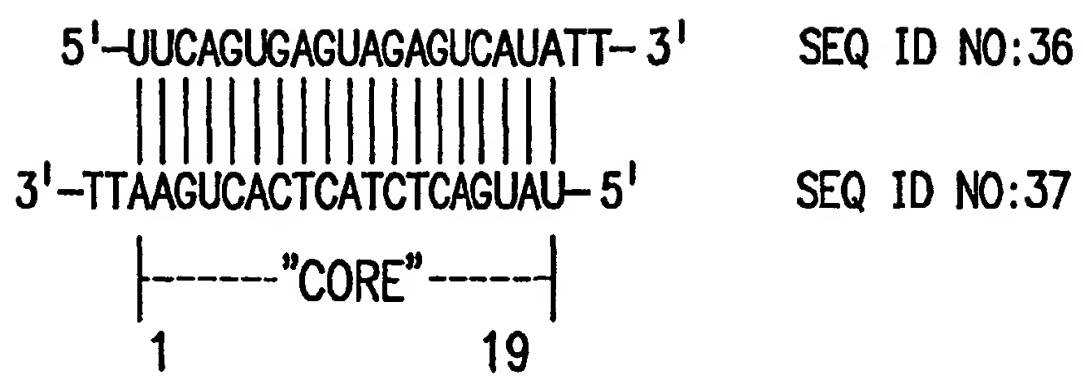


FIG.16

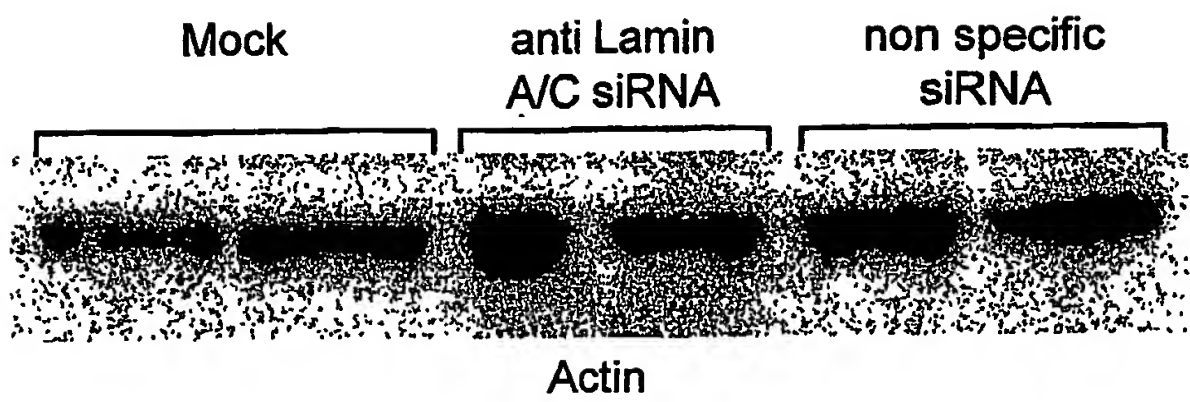


FIG.17A

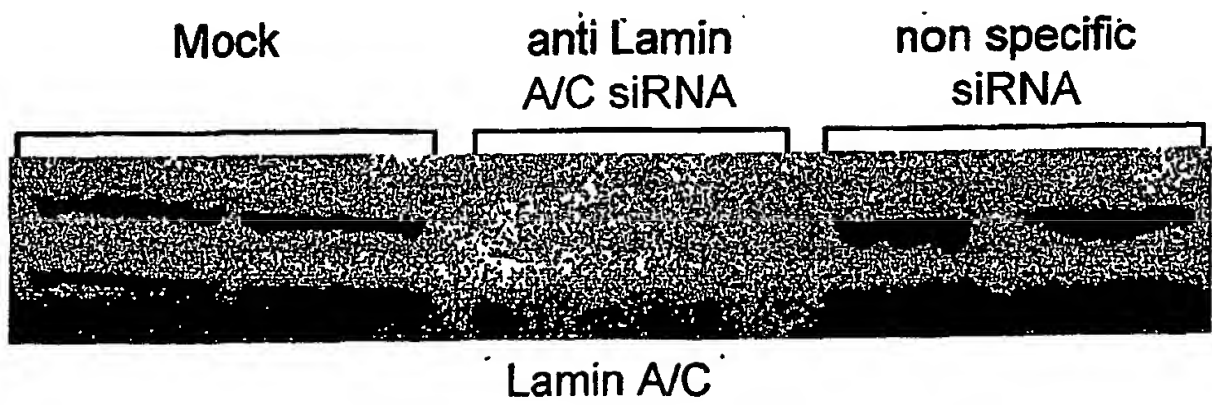


FIG.17B

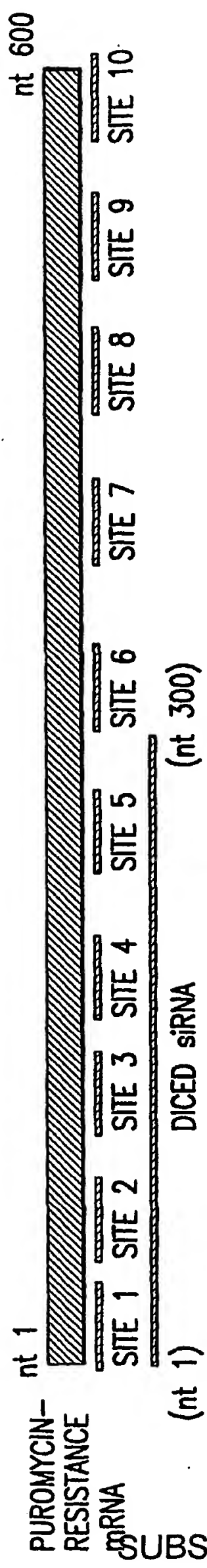


FIG.18A

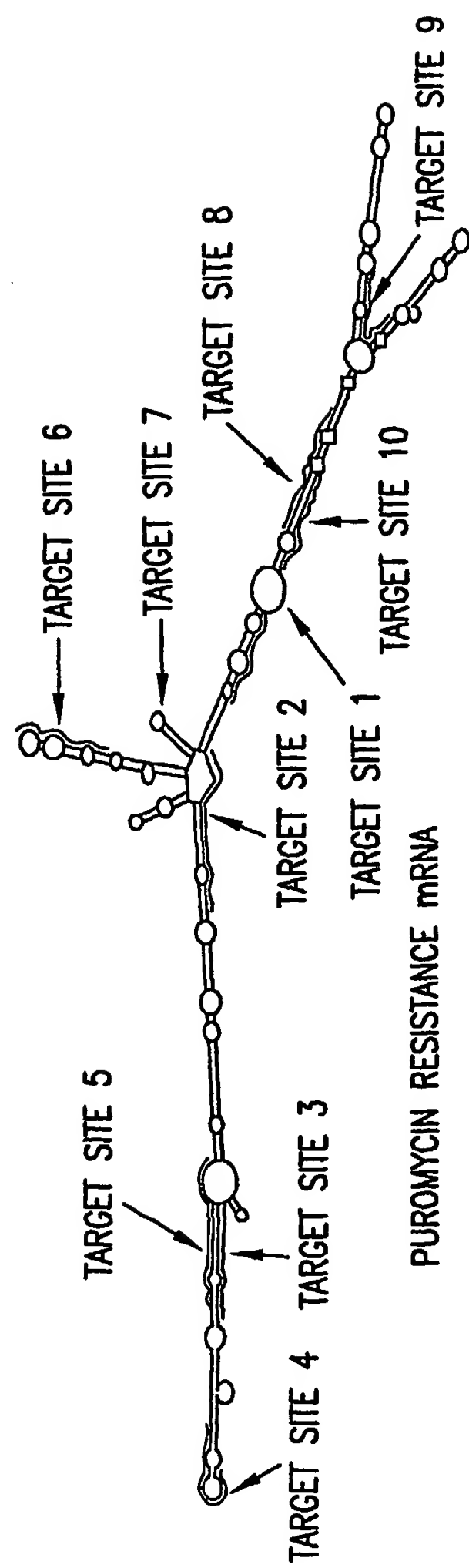


FIG.18B

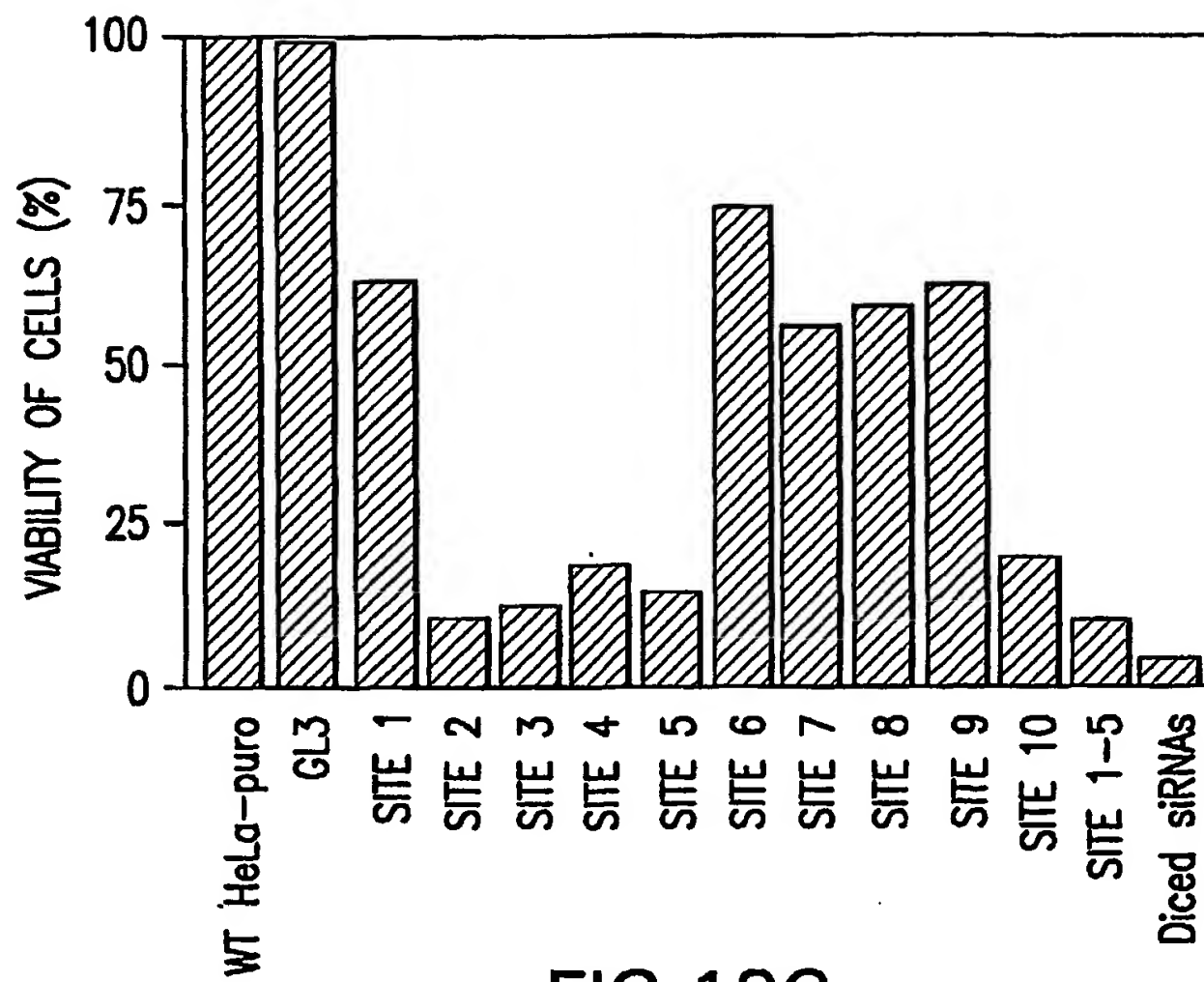
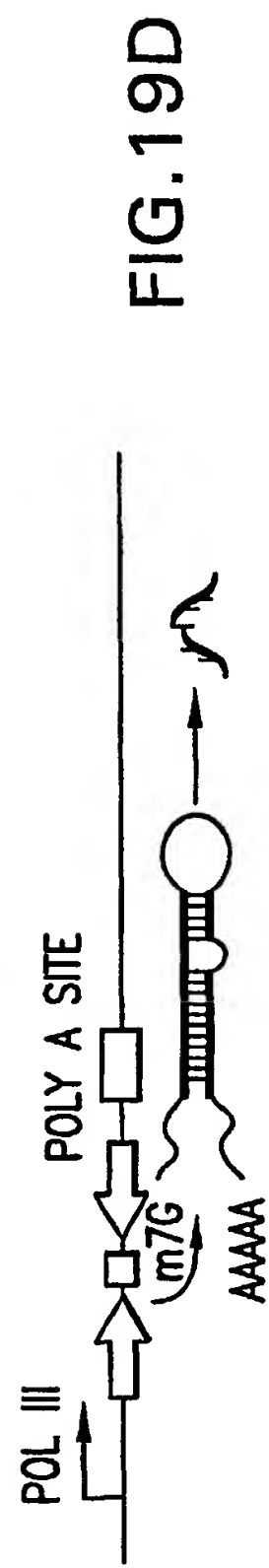
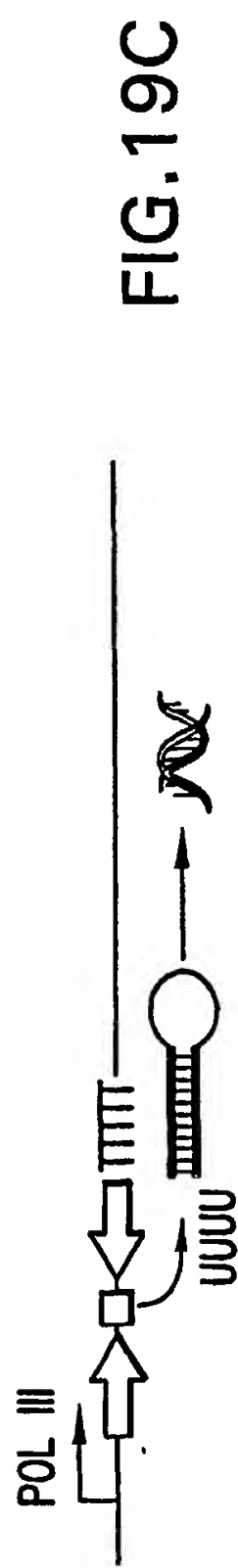
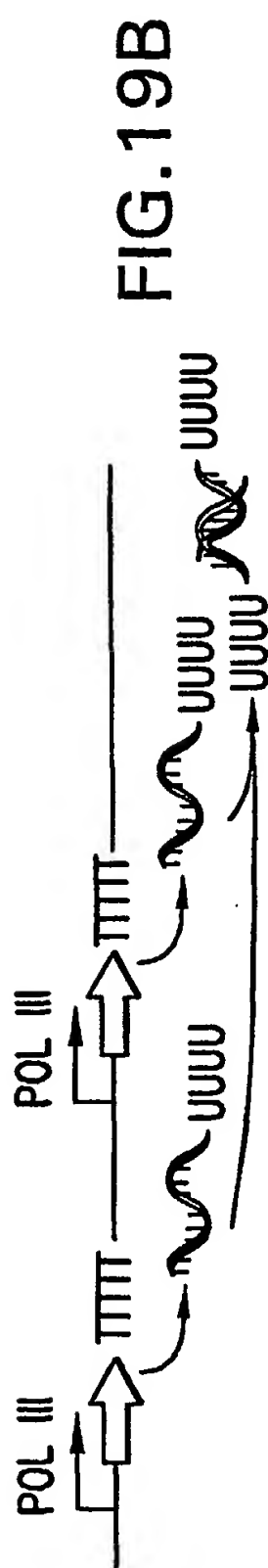
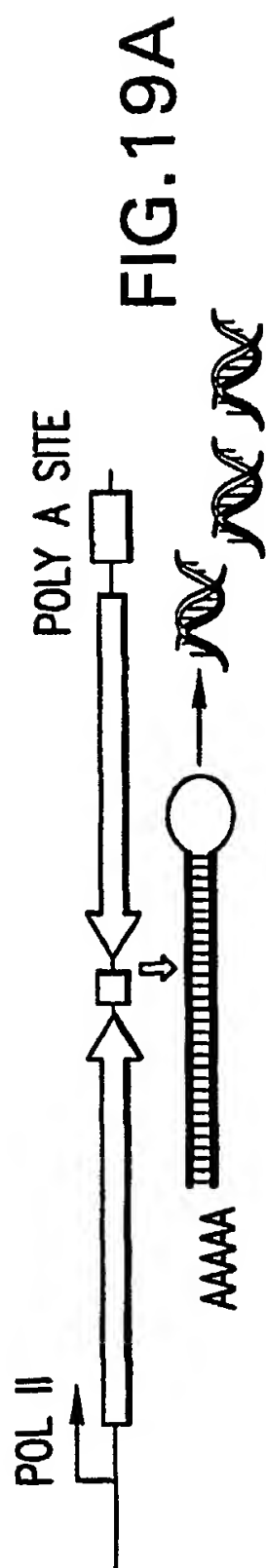


FIG. 18C





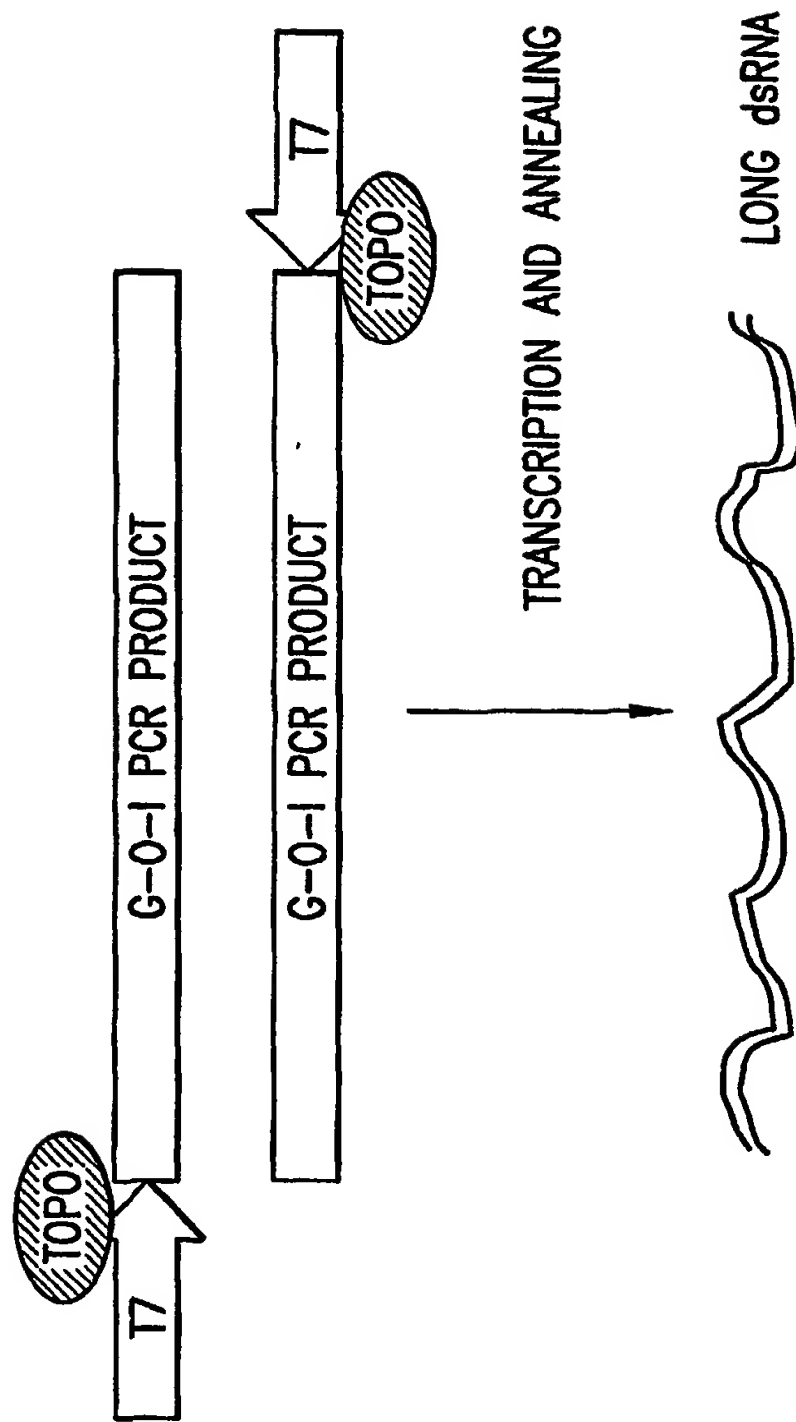


FIG.20

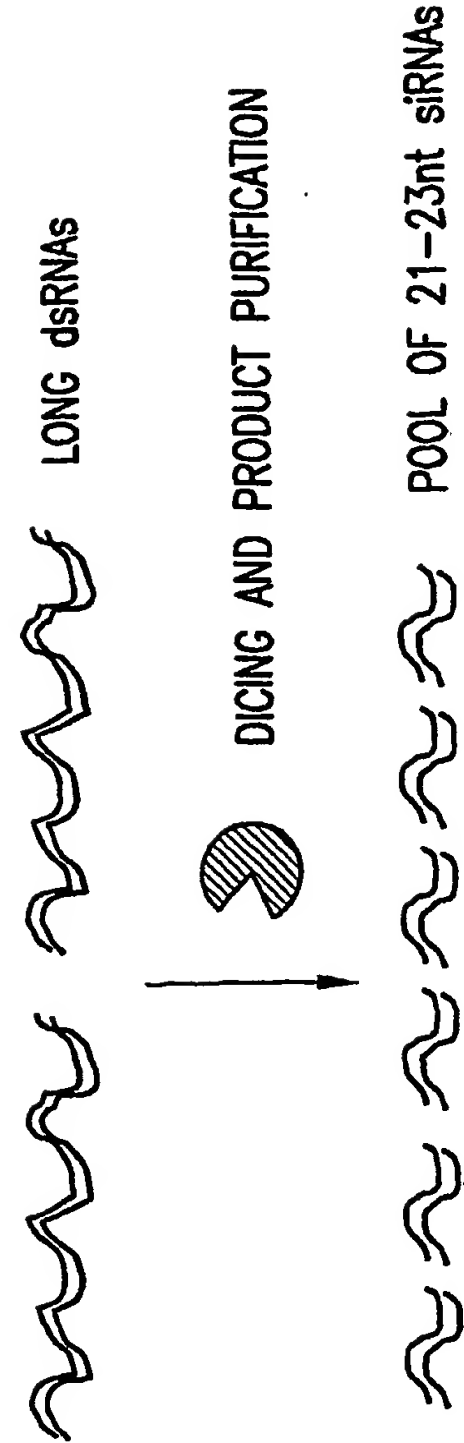


FIG.21

## Gel analysis



U PP P  
Diced  
Luciferase

U PP P  
Diced  
GFP

FIG.22A

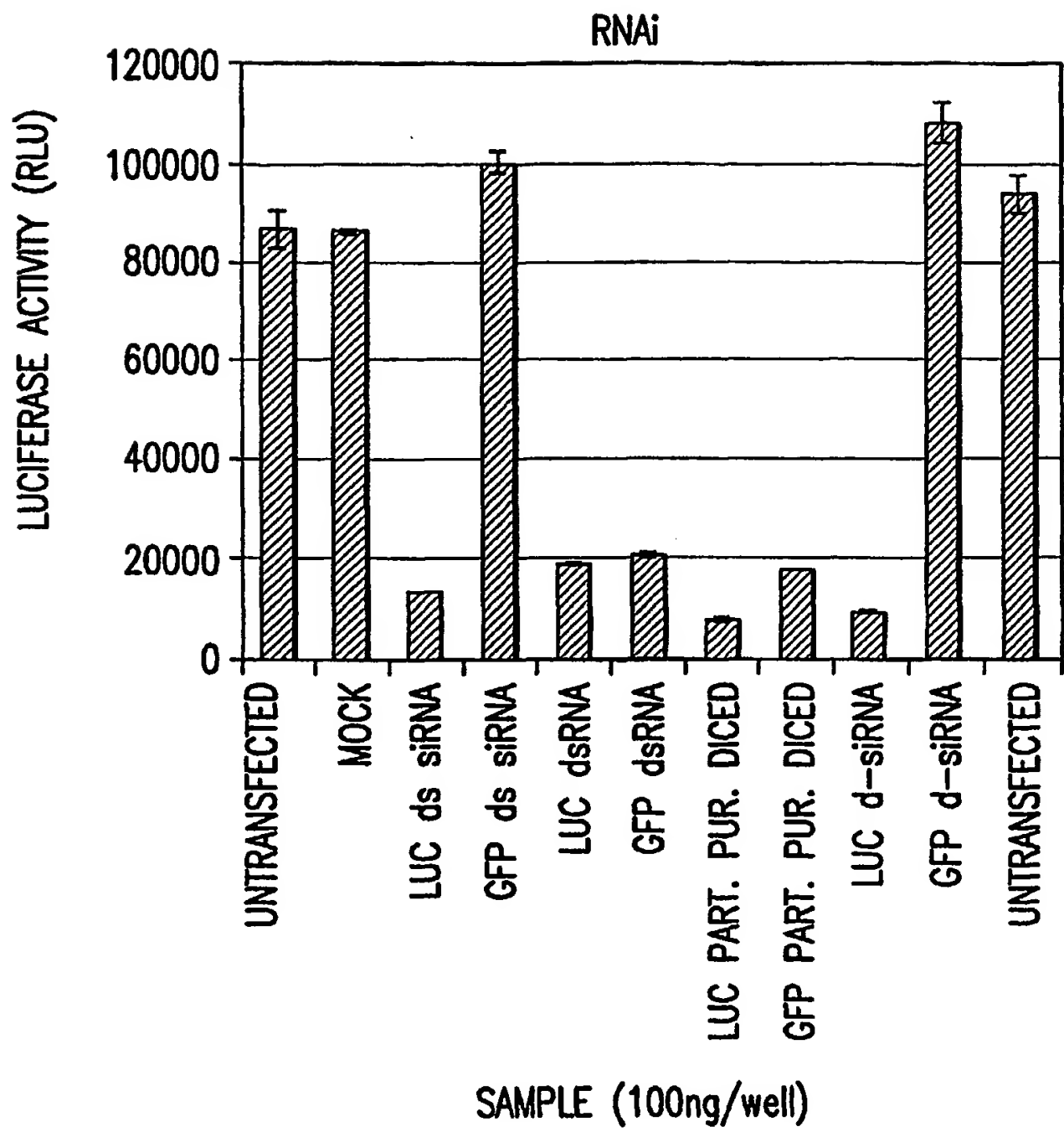


FIG.22B

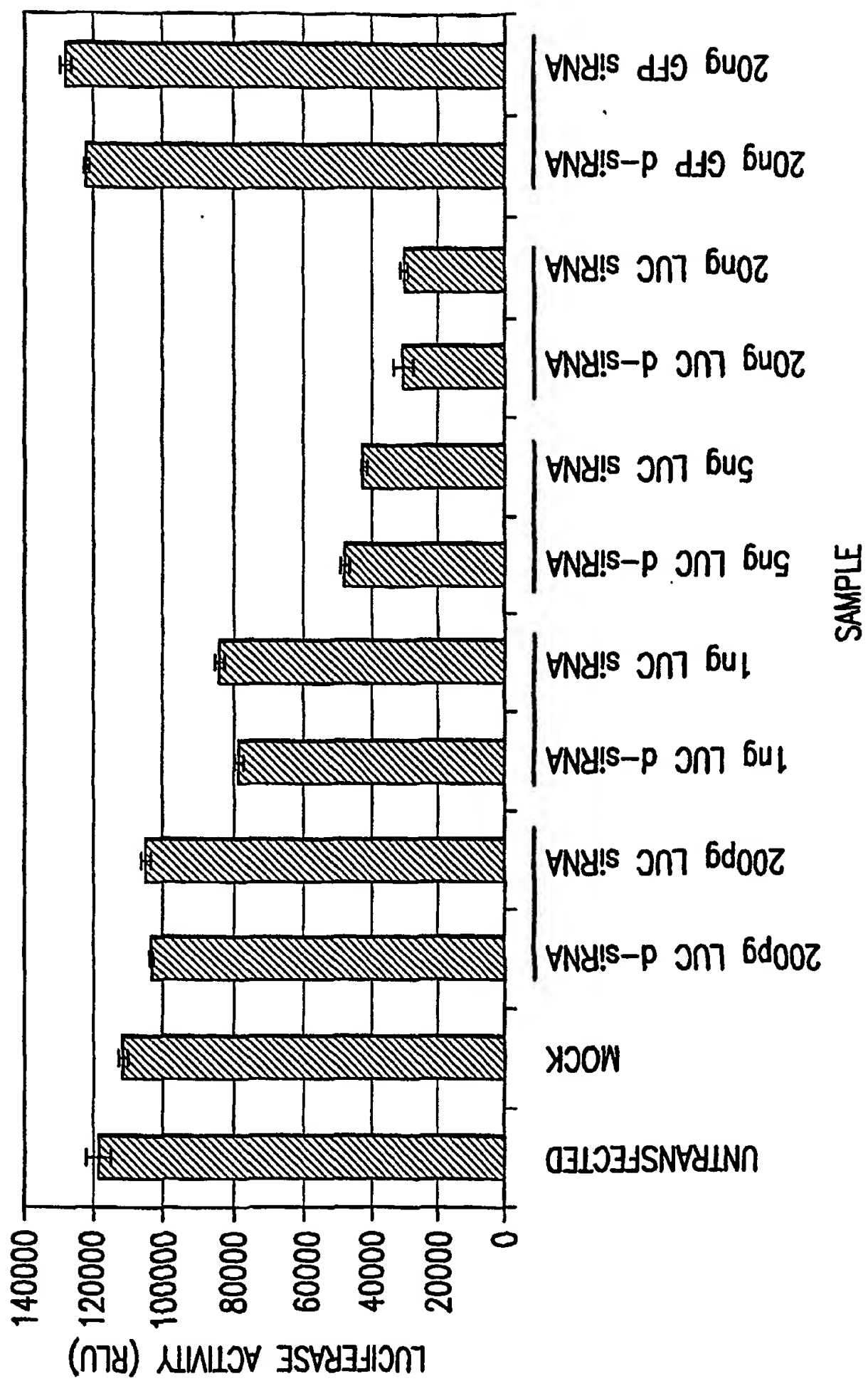


FIG. 23

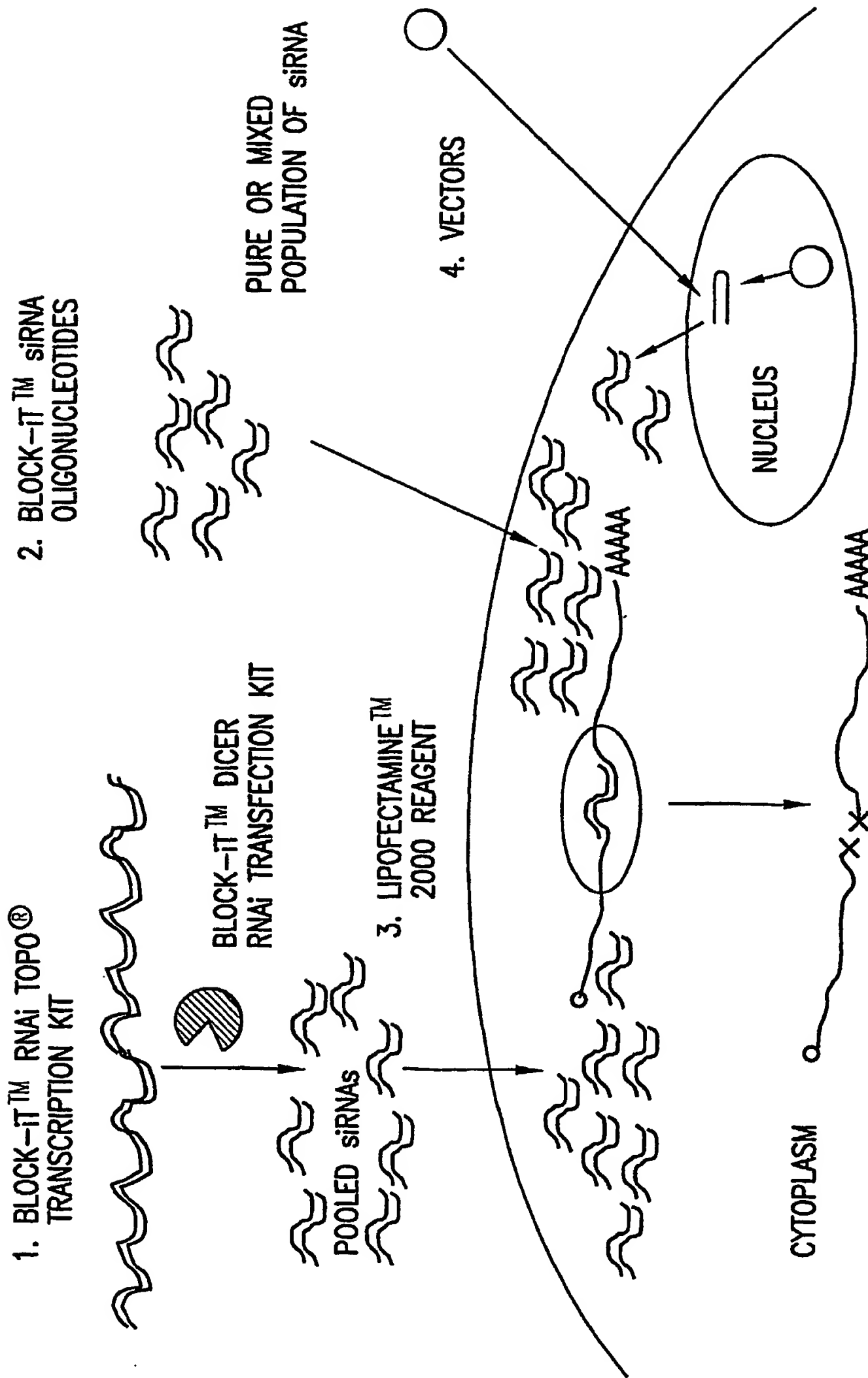


FIG.24





SUBSTITUTE SHEET (RULE 26)

- 36/49

FIG. 26

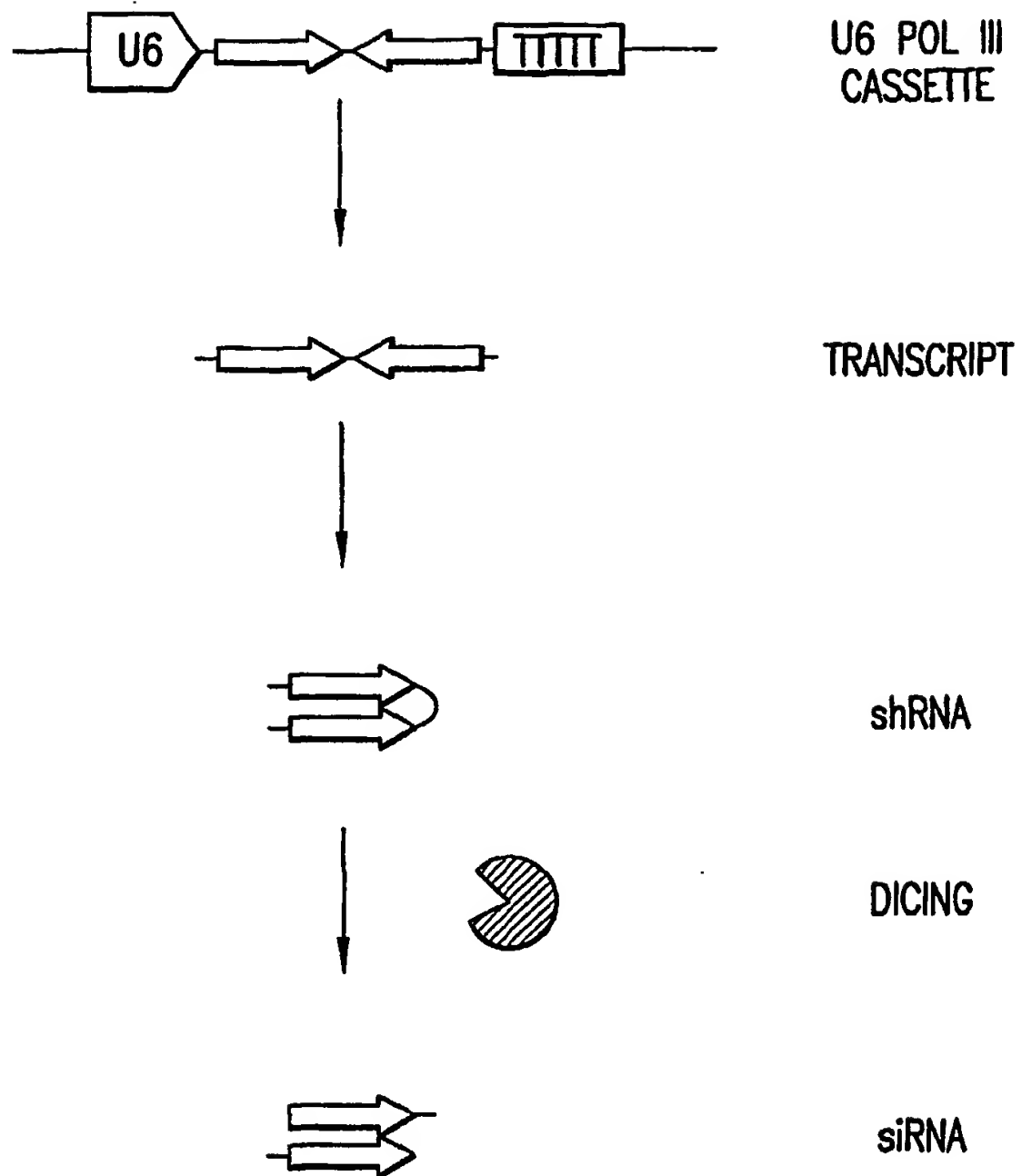


FIG.27

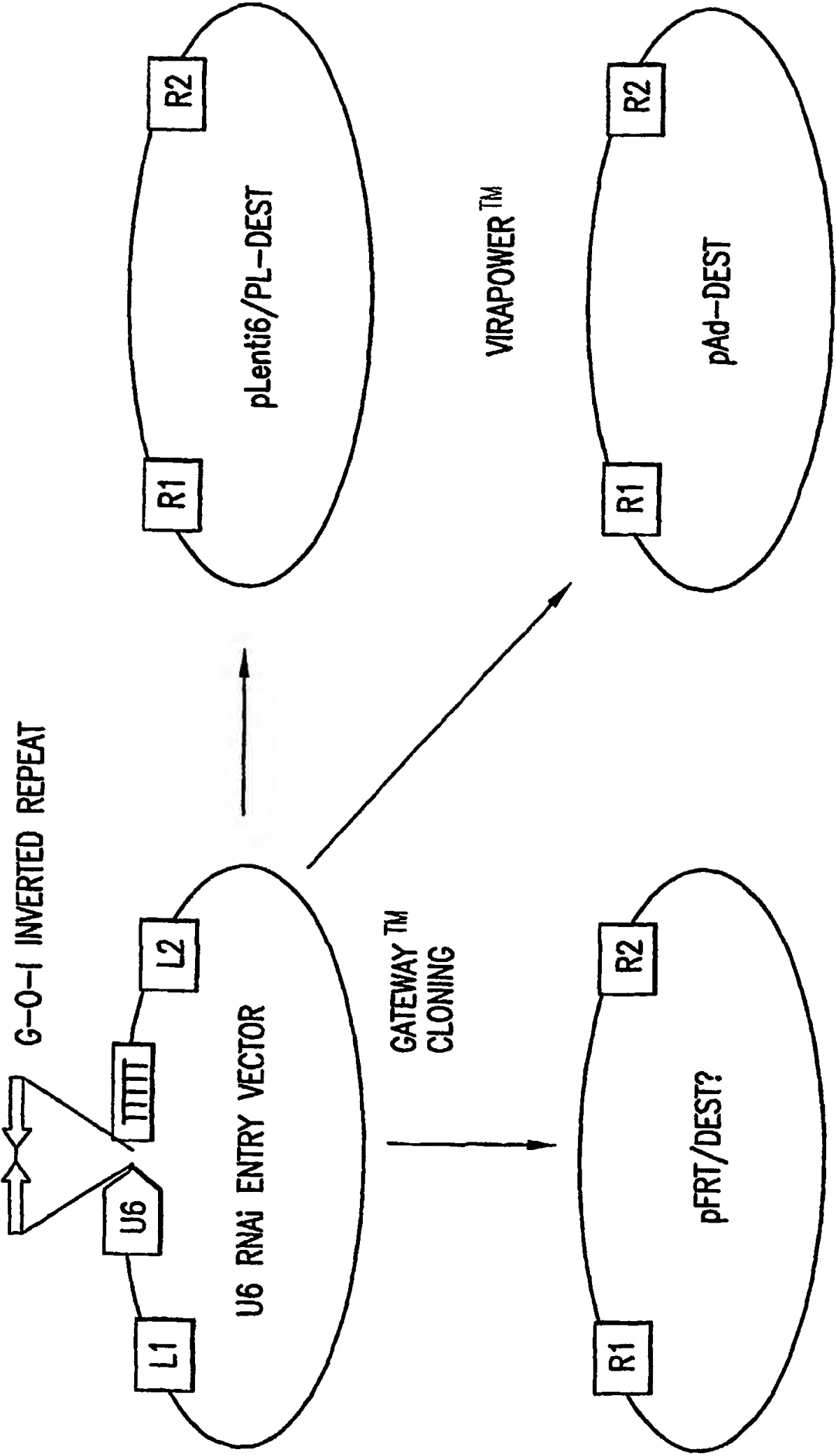


FIG.28

	TRANSIENT EXPRESSION		STABLE EXPRESSION			
	DIVIDING CELLS	NON-DIVIDING CELLS	DIVIDING CELLS	NEURONAL CELLS	DRUG OR GROWTH ARRESTED CELLS	CONTACT INHIBITED CELLS
ADENOVIRUS (DNA VIRUS)	YES	YES				
RETROVIRUS (RNA VIRUS)	YES		YES			
LENTIVIRUS (RNA VIRUS)	YES	YES	YES	YES	YES	YES

FIG.29

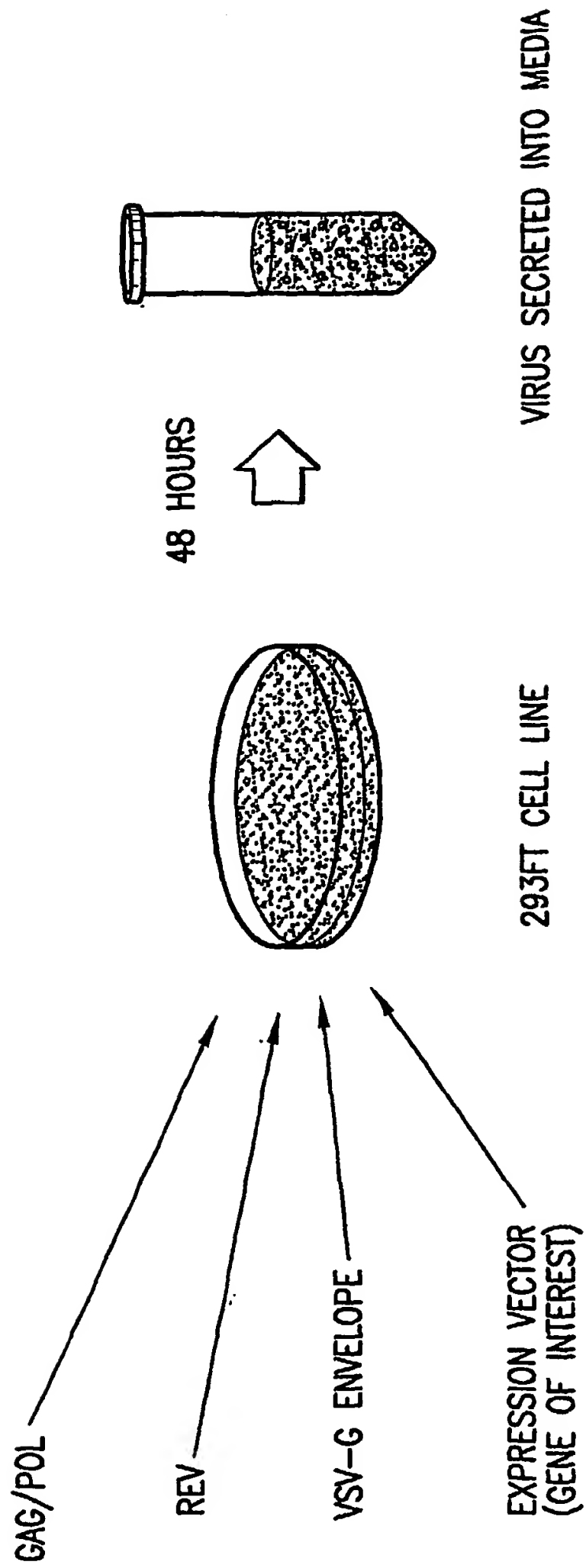


FIG.30

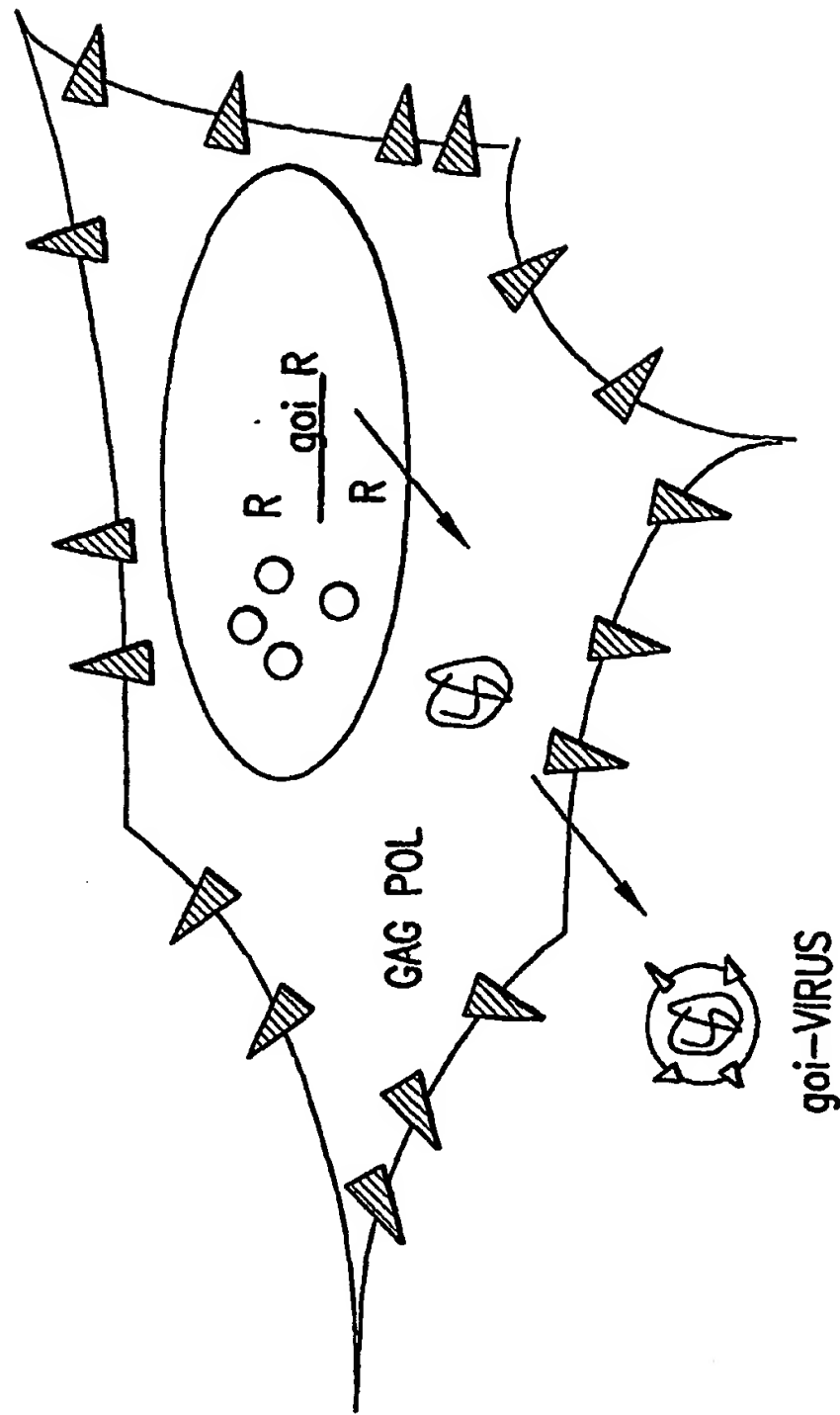
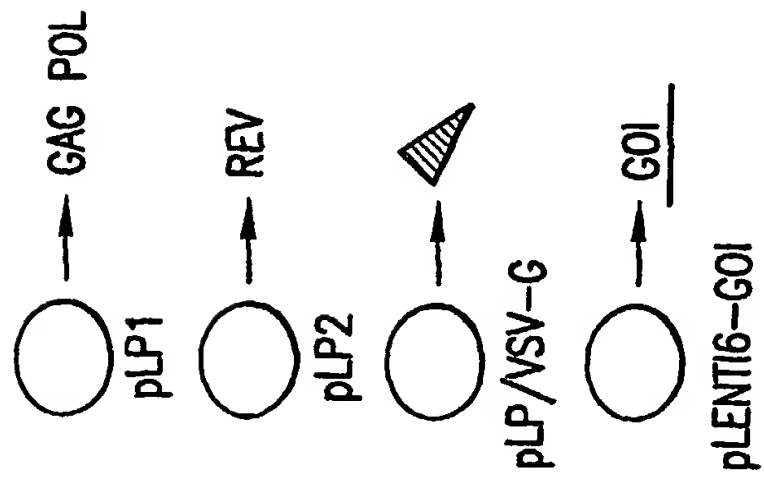


FIG.31



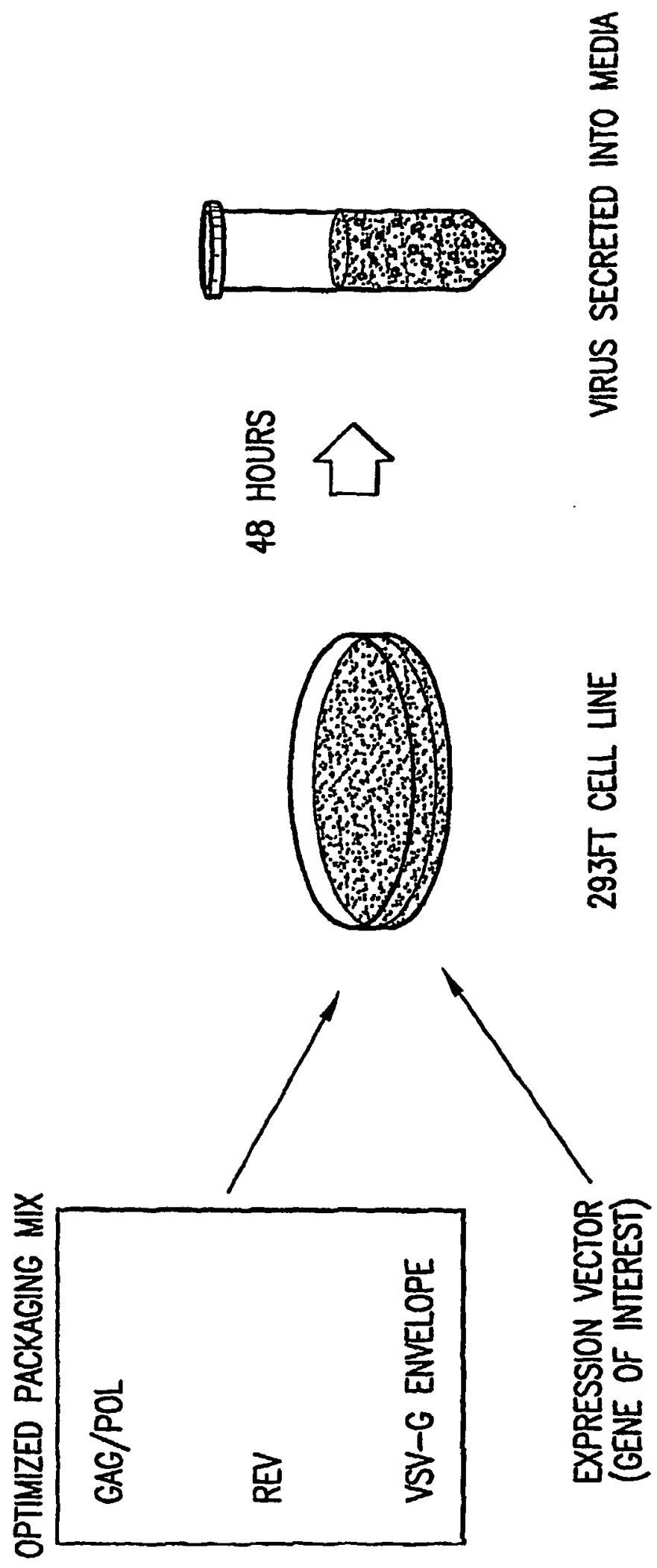
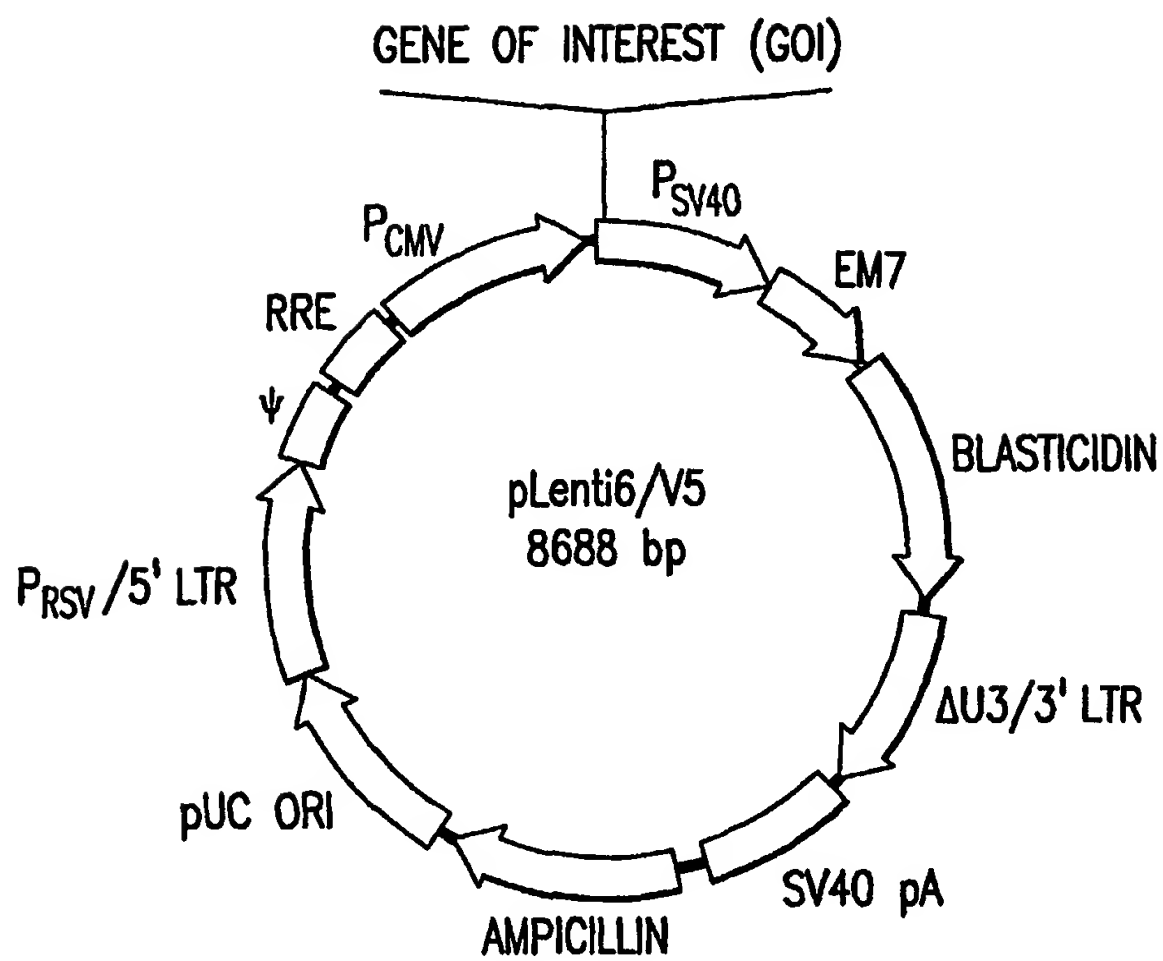


FIG.32

**FIG.33**



METHOD 1: DIRECTIONAL TOPO

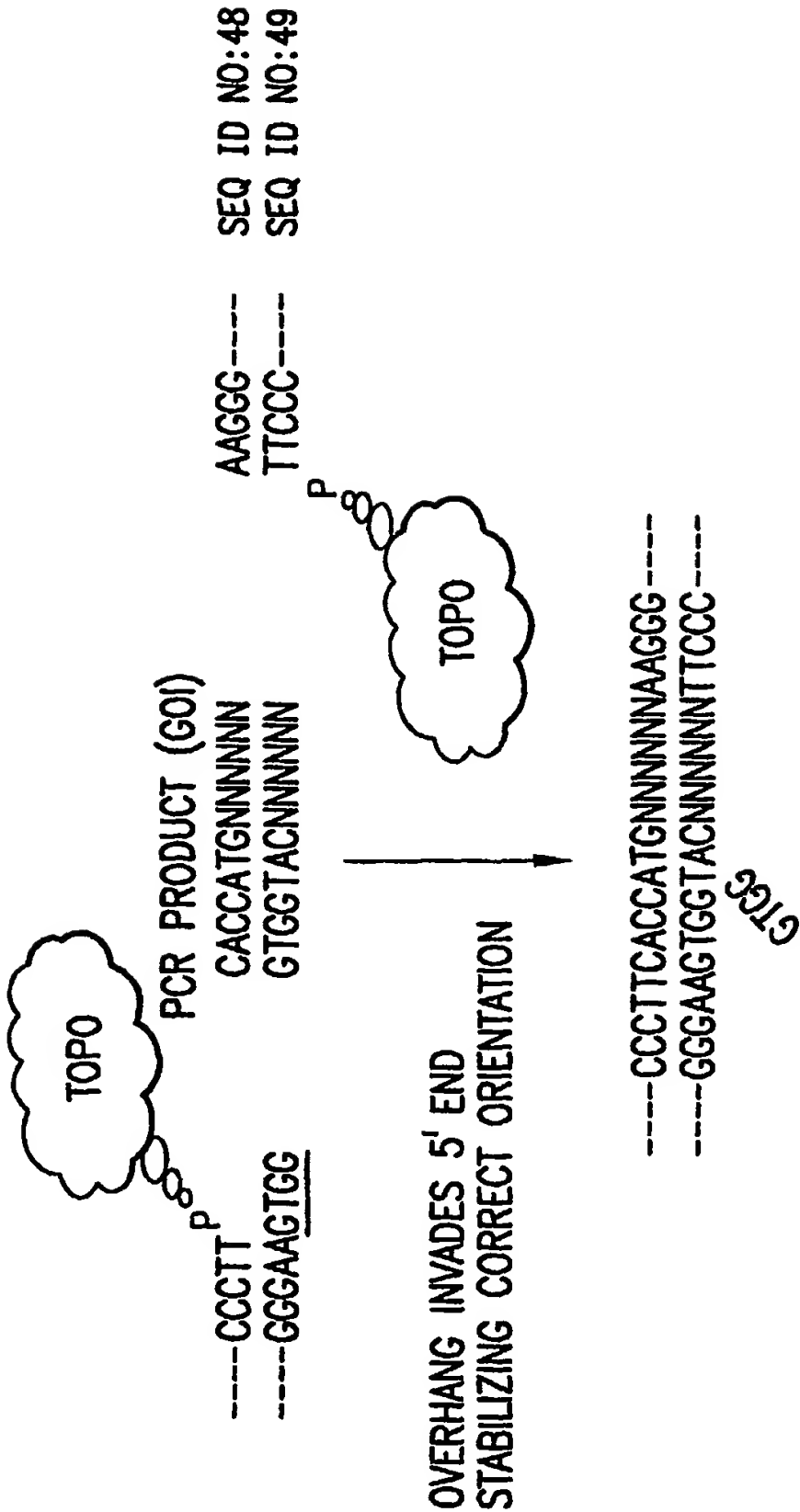


FIG. 34

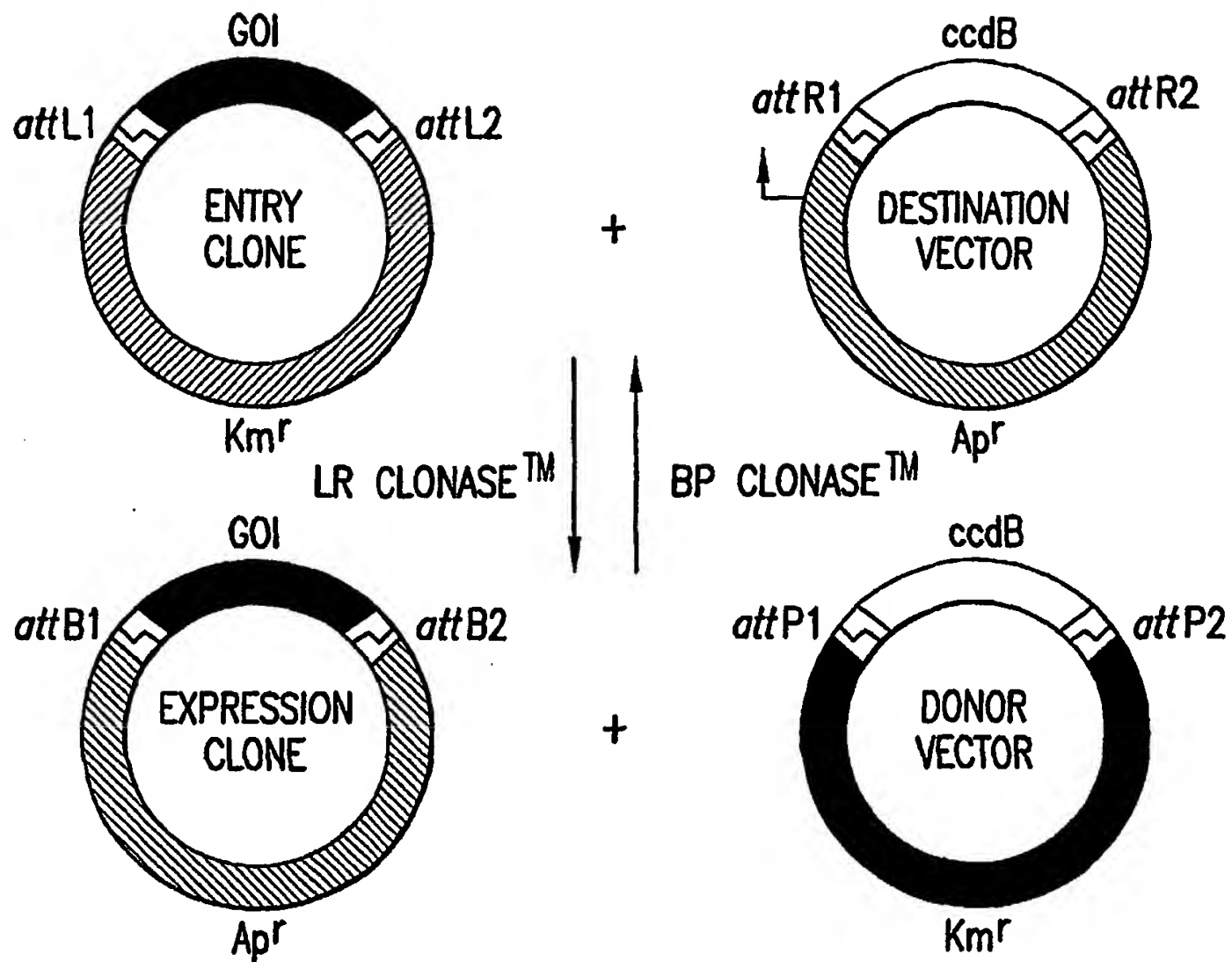
METHOD 2: GATEWAY™

FIG.35

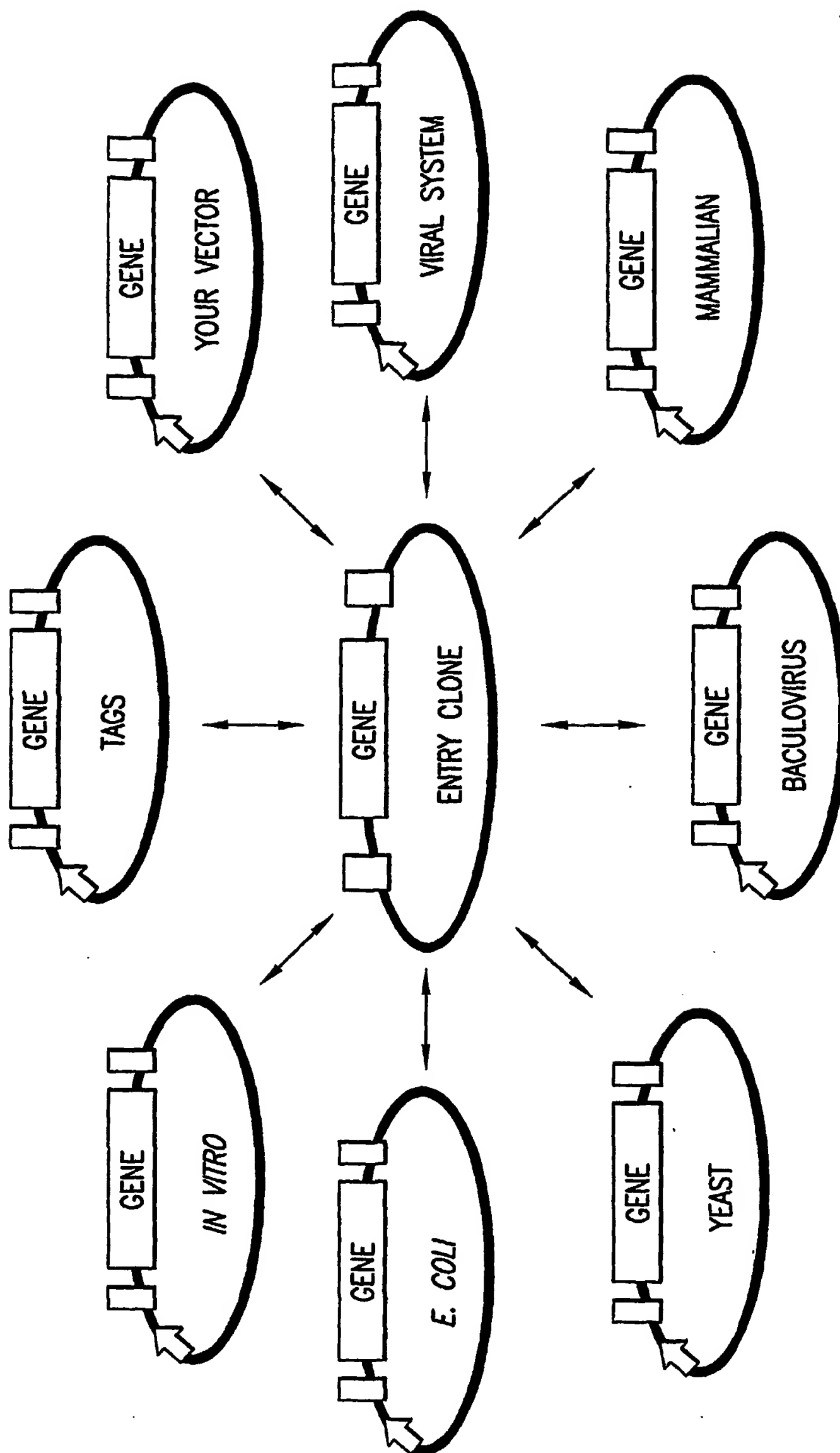


FIG.36

SUBSTITUTE SHEET (RULE 26)



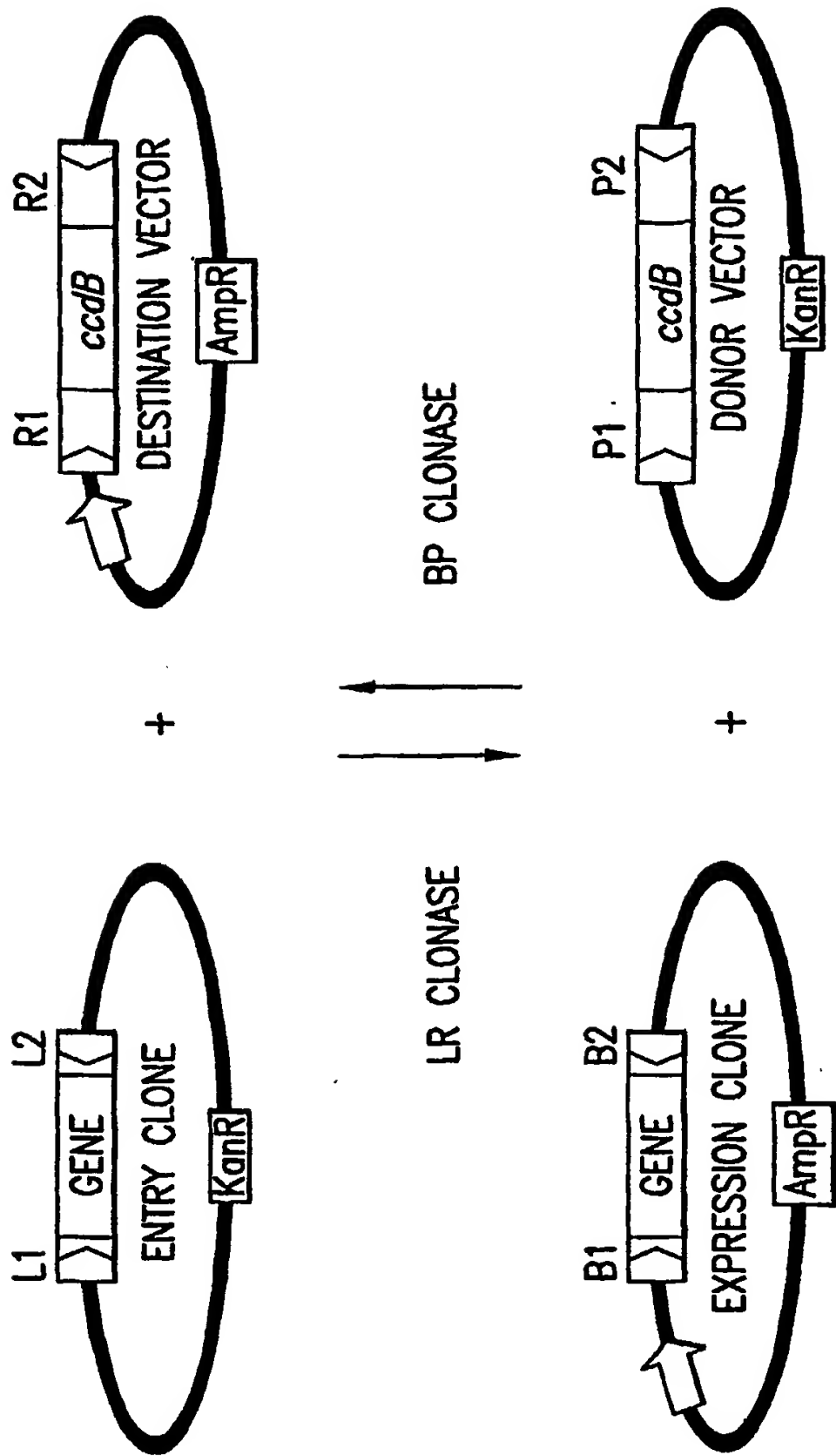


FIG.38

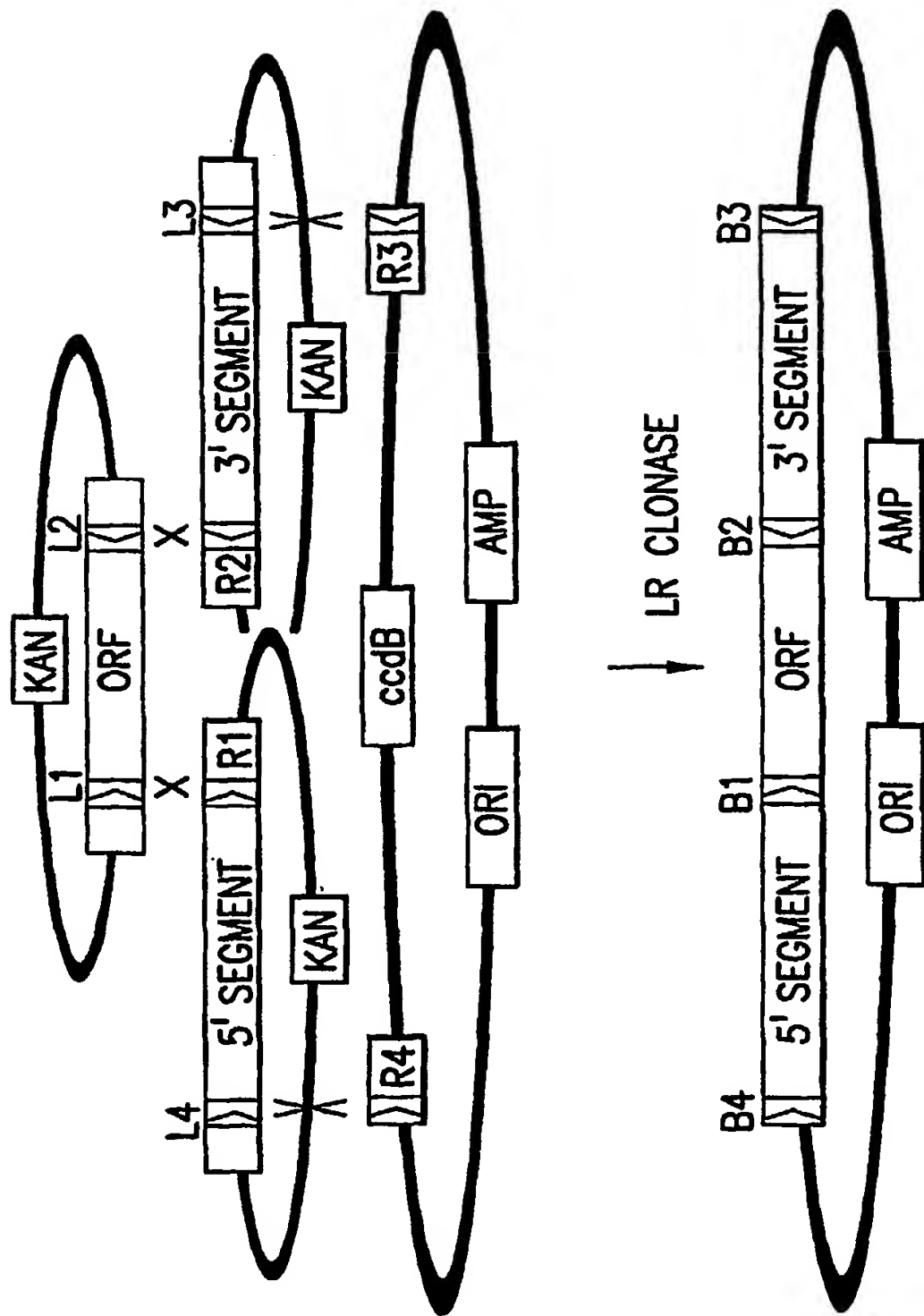


FIG.39

## SEQUENCE LISTING

<110> Invitrogen Corporation

Chesnut, Jon

Madden, Knut

Dudas, Miroslav

Leong, Louis

Harris, Adam

<120> METHODS AND COMPOSITIONS FOR PERFORMING SEAMLESS CLONING

<130> 0942.581PC02

<150> 60/493,322

<151> 2003-08-08

<160> 51

<170> PatentIn version 3.0

<210> 1

<211> 3519

<212> DNA

<213> Artificial

<220>

<223> pENTRU6 Vector

<400> 1

ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga 60

taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga 120

gcgccaata cgcaaaccgc ctctccccgc gcgttggccg attcattaat gcagctggca 180

cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaata cgcgtagcgc	240
tagccaggaa gagtttgtag aaacgcaaaa aggccatccg tcaggatggc cttctgctta	300
gtttgatgcc tggcagttta tggcgggcgt cctgcccgc accctccggg ccgttgcttc	360
acaacgttca aatccgctcc cggcggattt gtcctactca ggagagcgtt caccgacaaa	420
caacagataa aacgaaaggc ccagtcttcc gactgagcct ttcgttttat ttgatgcctg	480
gcagttccct actctcgcg taacgctagc atggatgttt tcccagtcac gacgttgtaa	540
aacgacggcc agtcttaagc tcggggccca aataatgatt ttattttgac tgatagtga	600
ctgttcgttg caacaaattg atgagcaatg cttttttata atgccaaactt tgtacaaaaa	660
agcaggcttt aaaggaacca attcagtcga ctggatccgg taccaaggtc gggcaggaag	720
agggcctatt tcccatgatt ccttcatatt tgcataatcg atacaaggct gttagagaga	780
taattagaat taatttgact gtaaacacaa agatattagt acaaaatagc tgacgtagaa	840
agtaataatt tcttgggtag tttgcagttt taaaattatg ttttaaaatg gactatcata	900
tgcttaccgt aacttgaaag tatttcgatt tcttggtttt atatatcttg tggaaaggac	960
gaaacaccgg agaccgcggc cgctggatcc ggcttactaa aagccagata acagtatgcg	1020
tatttgcgcg ctgatttttg cggataaaga atatatactg atatgtatac ccgaagtatg	1080
tcaaaaagag gtgtgctatg aagcagcgta ttacagtga agttgacagc gacagctatc	1140
agttgctcaa ggcataatg atgtcaatat ctccggtctg gtaagcacia ccatgcagaa	1200
tgaagcccgt cgtctgcgtg ccgaacgctg gaaagcggaa aatcaggaag ggatggctga	1260
ggtcgcccgg tttattgaaa tgaacggctc ttttgctgac gagaacaggg actggtgaaa	1320
tgcaagttta ggtttacacc tataaaagag agagccgtta tcgtctgttt gtggatgtac	1380
agagtgatat tattgacacg cccgggcgac ggatgggtgat cccctggcc agtgcacgtc	1440
tgctgtcaga taaagtctcc cgtgaacttt acccggtggg gcatatcggg gatgaaagct	1500
ggcgcagatg gaccaccgat atggccagtg tgccggtctc cgttatcggg gaagaagtgg	1560
ctgatctcag ccaccgcgaa aatgacatca aaaacgccat taacctgatg ttctggggaa	1620
tataaggctc catttttttt ctagaccag ctttcttgta caaagttggc attataagaa	1680
agcattgctt atcaatttgt tgcaacgaac aggtcactat cagtcaaaat aaaatcatta	1740
tttgccatcc agctgatatc ccctatagtg agtcgtatta catggtcata gctgtttcct	1800
ggcagctctg gcccggtgtc caaaatctct gatgttacat tgcacaagat aaaaatatat	1860
catcatgaac aataaaactg tctgcttaca taaacagtaa tacaaggggt gttatgagcc	1920
atattcaacg ggaaacgtcg aggccgcgat taaattccaa catggatgct gatttatatg	1980
ggtataaatg ggctcgcgat aatgtcgggc aatcaggtgc gacaatctat cgcttgatg	2040



```

ggaagcccga tgcgccagag ttgtttctga aacatggcaa aggtagcgtt gccaatgatg 2100
ttacagatga gatggtcaga ctaaactggc tgacggaatt tatgcctctt ccgaccatca 2160
agcattttat ccgtactcct gatgatgcat ggttactcac cactgcgac cccggaaaaa 2220
cagcattcca ggtattagaa gaatatcctg attcaggtga aaatattgtt gatgcgctgg 2280
cagtgttcct gcgccggttg cattcgattc ctgtttgtaa ttgtcctttt aacagcgatc 2340
gcgtatttcg tctcgctcag gcgcaatcac gaatgaataa cggtttggtt gatgcgagtg 2400
at ttgatgac gagcgtaatg gctggcctgt tgaacaagtc tggaaagaaa tgcataaact 2460
tttgccattc tcaccggatt cagtcgtcac tcatgggtgat ttctcacttg ataaccttat 2520
ttttgacgag gggaaattaa taggttgat tgatgttgga cgagtcggaa tcgcagaccg 2580
ataccaggat cttgccatcc tatggaactg cctcggtgag ttttctcctt cattacagaa 2640
acggcttttt caaaaatatg gtattgataa tcctgatatg aataaattgc agtttcattt 2700
gatgctcgat gagtttttct aatcagaatt ggttaattgg ttgtaacact ggcagagcat 2760
tacgctgact tgacgggacg gcgcaagctc atgaccaaaa tcccttaacg tgagttacgc 2820
gtcgttccac tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt 2880
ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cggtggtttg 2940
tttgccggat caagagctac caactctttt tccgaaggta actggcttca gcagagcgca 3000
gataccaaat actgtccttc tagtgtagcc gtagttaggc caccacttca agaactctgt 3060
agcaccgcct acatacctcg ctctgcta at cctgttacca gtggctgctg ccagtggcga 3120
taagtcgtgt cttaccgggt tggactcaag acgatagtta ccggataagg cgcagcggtc 3180
gggctgaacg ggggggttcgt gcacacagcc cagcttgag cgaacgacct acaccgaact 3240
gagataccta cagcgtgagc attgagaaag cgccacgctt cccgaaggga gaaaggcgga 3300
caggtatccg gtaagcggca gggtcggaac aggagagcgc acgagggagc ttccaggggg 3360
aaacgcctgg tatctttata gtcctgtcgg gtttcgccac ctctgacttg agcgtcgatt 3420
tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac gccagcaacg cggccttttt 3480
acggttcctg gccttttgct ggccttttgc tcacatgtt 3519

```

<210> 2

<211> 8688

<212> DNA

<213> Artificial

&lt;220&gt;

&lt;223&gt; pLenti6/V5-DEST

&lt;400&gt; 2

aatgtagtct tatgcaatac tctttagtagtc ttgcaacatg gtaacgatga gtttagcaaca	60
tgccttataca ggagagaaaa agcaccgtgc atgccgattg gtggaagtaa ggtgggtacga	120
tcgtgcctta ttaggaaggc aacagacggg tctgacatgg attggacgaa ccaactgaatt	180
gccgcattgc agagatatattg tattttaagtg cctagctcga tacataaacg ggtctctctg	240
gtttagaccag atctgagcct gggagctctc tggctaacta ggaacccac tgcttaagcc	300
tcaataaagc ttgccttgag tgcttcaagt agtgtgtgcc cgtctgttgt gtgactctgg	360
taactagaga tccctcagac ccttttagtc agtgtggaaa atctctagca gtggcgccc	420
aacagggact tgaaagcgaa agggaaacca gagtagctct ctcgacgcag gactcggctt	480
gctgaagcgc gcacggcaag aggcgagggg cggcgactgg tgagtacgcc aaaaattttg	540
actagcggag gctagaagga gagagatggg tgcgagagcg tcagtattaa gcgggggaga	600
attagatcgc gatgggaaaa aattcgggta aggccagggg gaaagaaaa atataaatta	660
aaacatatag tatgggcaag caggagacta gaacgattcg cagttaatcc tggcctgtta	720
gaaacatcag aaggctgtag acaaatactg ggacagctac aaccatccct tcagacagga	780
tcagaagaac ttagatcatt atataatata gtagcaacc tctatttgtgt gcatcaaagg	840
atagagataa aagacaccaa ggaagcttta gacaagatag aggaagagca aaacaaaagt	900
aagaccaccg cacagcaagc ggccgctgat cttcagacct ggaggaggag atatgaggga	960
caattggaga agtgaattat ataaatataa agtagtaaaa attgaaccat taggagtagc	1020
accaccaag gcaaagagaa gagtgggtgca gagagaaaa agagcagtgg gaataggagc	1080
tttgttcctt gggttccttg gagcagcagg aagcactatg ggcgcagcgt caatgacgct	1140
gacggtacag gccagacaat tattgtcttg tatagtgcag cagcagaaca atttgctgag	1200
ggctattgag gcgcaacagc atctgttgca actcacagtc tggggcatca agcagctcca	1260
ggcaagaatc ctggctgtgg aaagatacct aaaggatcaa cagctcctgg ggatttgggg	1320
ttgctctgga aaactcattt gcaccactgc tgtgccttg aatgctagtt ggagtaataa	1380
atctctggaa cagatttgga atcacacgac ctggatggag tgggacagag aaattaacaa	1440
ttacacaagc ttaatacact ccttaattga agaatcgcaa aaccagcaag aaaagaatga	1500
acaagaatta ttggaattag ataaatgggc aagtttgtgg aattgggtta acataacaaa	1560
ttggctgtgg tatataaaat tattcataat gatagtagga ggcttggtag gtttaagaat	1620

agtttttgct	gtactttcta	tagtgaatag	agttaggcag	ggatattcac	cattatcggt	1680
tcagacccac	ctcccaaccc	cgagggggacc	cgacaggccc	gaaggaatag	aagaagaagg	1740
tggagagaga	gacagagaca	gatccattcg	attagtgaac	ggatctcgac	ggtatcgata	1800
agcttgggag	ttccgcgtta	cataacttac	ggtaaattggc	ccgcctggct	gaccgcccaa	1860
cgacccccgc	ccattgacgt	caataatgac	gtatgttccc	atagtaacgc	caataggggac	1920
tttccattga	cgtcaatggg	tggagtattt	acggtaaact	gccacttgg	cagtacatca	1980
agtgtatcat	atgccaaagta	cgccccctat	tgacgtcaat	gacggtaaata	ggcccgctg	2040
gcattatgcc	cagtacatga	ccttatggga	ctttcctact	tggcagtaca	tctacgtatt	2100
agtcacgct	attaccatgg	tgatgcgggt	ttggcagtac	atcaatgggc	gtggatagcg	2160
gtttgactca	cggggatttc	caagtctcca	ccccattgac	gtcaatggga	gtttgttttg	2220
gcacaaaaat	caacgggact	ttccaaaatg	tcgtaacaac	tccgccccat	tgacgcaaat	2280
gggcggtagg	cgtgtacggg	gggaggtcta	tataagcaga	gctcgtttag	tgaaccgtca	2340
gatcgctgg	agacgccatc	cacgctgttt	tgacctccat	agaagacacc	gactctagag	2400
gatccactag	tccagtgtgg	tggaaattctg	cagatatcaa	caagtttgta	caaaaaagct	2460
gaacgagaaa	cgtaaaatga	tataaatatc	aatatattaa	attagatttt	gcataaaaaa	2520
cagactacat	aatactgtaa	aacacaacat	atccagtcac	tatggcggcc	gcattaggca	2580
ccccaggctt	tacactttat	gcttccggct	cgtataatgt	gtggattttg	agttaggatc	2640
cggcgagatt	ttcaggagct	aaggaagcta	aatggagaa	aaaaatcact	ggatatacca	2700
ccgttgatat	atcccaatgg	catcgtaaag	aacattttga	ggcatttcag	tcagttgctc	2760
aatgtaccta	taaccagacc	gttcagctgg	atattacggc	ctttttaaag	accgtaaaga	2820
aaaataagca	caagttttat	ccggccttta	ttcacattct	tgcccgctg	atgaatgctc	2880
atccggaatt	ccgtatggca	atgaaagacg	gtgagctggg	gatatgggat	agtgttcacc	2940
cttgttacac	cgttttccat	gagcaaactg	aaacgttttc	atcgctctgg	agtgaatacc	3000
acgacgattt	ccggcagttt	ctacacatat	attcgcaaga	tgtggcgtgt	tacggtgaaa	3060
acctggccta	tttccctaaa	gggtttattg	agaatatgtt	tttcgtctca	gccaatccct	3120
gggtgagttt	caccagtttt	gatttaaacg	tggccaatat	ggacaacttc	ttcgcccccg	3180
ttttcaccat	gggcaaatat	tatacgcaag	gcgacaagg	gctgatgccg	ctggcgattc	3240
aggttcatca	tgccgtctgt	gatggcttcc	atgtcggcag	aatgcttaata	gaattacaac	3300
agtactgcga	tgagtggcag	ggcggggcgt	aaagatctgg	atccggctta	ctaaaagcca	3360
gataacagta	tcggtatttg	cgcgctgatt	tttgcggtat	aagaatatat	actgatatgt	3420
ataccggaag	tatgtcaaaa	agaggtgtgc	tatgaagcag	cgtattacag	tgacagttga	3480

cagcgacagc	tatcagttgc	tcaaggcata	tatgatgtca	atatctccgg	tctggtaagc	3540
acaaccatgc	agaatgaagc	ccgtcgtctg	cgtgccgaac	gctggaaagc	ggaaaatcag	3600
gaagggatgg	ctgaggtcgc	ccggtttatt	gaaatgaacg	gctcttttgc	tgacgagaac	3660
agggactggc	gaaatgcagt	ttaaggttta	cacctataaa	agagagagcc	gttatcgtct	3720
gtttgtggat	gtacagagtg	atattattga	cacgcccggg	cgacggatgg	tgatccccct	3780
ggccagtgc	cgtctgctgt	cagataaagt	ctcccgtgaa	ctttaccogg	tggtgcatat	3840
cggggatgaa	agctggcgca	tgatgaccac	cgatatggcc	agtgtgccgg	tctccgttat	3900
cggggaagaa	gtggctgac	tcagccaccg	cgaaaatgac	atcaaaaacg	ccattaacct	3960
gatgttctgg	ggaatataaa	tgtcaggctc	cgttatacac	agccagtctg	caggtcgacc	4020
atagtactg	gatatgttgt	gttttacagt	attatgtagt	ctgtttttta	tgcaaaatct	4080
aatttaatat	attgatattt	atatcatttt	acgtttctcg	ttcagctttc	ttgtacaaag	4140
tggttgatat	ccagcacagt	ggcggccgct	cgagtctaga	gggcccgcgg	ttcgaaggta	4200
agcctatccc	taaccctctc	ctcggctctg	attctacgcg	taccggttag	taatgagttt	4260
ggaattaatt	ctgtggaatg	tgtgtcagtt	aggggtgtga	aagtccccag	gctccccagg	4320
caggcagaag	tatgcaaagc	atgcatctca	attagtcagc	aaccagggtg	ggaaagtccc	4380
caggctcccc	agcaggcaga	agtatgcaaa	gcatgcatct	caattagtca	gcaaccatag	4440
tcccgcacct	aactccgccc	atcccgcacc	taactccgcc	cagttccgcc	cattctccgc	4500
cccatggctg	actaattttt	tttatttatg	cagaggccga	ggccgcctct	gcctctgagc	4560
tattccagaa	gtagtgagga	ggcttttttg	gaggcctagg	cttttgcaaa	aagctcccgg	4620
gagcttgat	atccattttc	ggatctgac	agcacgtgtt	gacaattaat	catcggcata	4680
gtatatcggc	atagtataat	acgacaaggt	gaggaactaa	accatggcca	agcctttgtc	4740
tcaagaagaa	tccaccctca	ttgaaagagc	aacggctaca	atcaacagca	tccccatctc	4800
tgaagactac	agcgtcgcca	gcgcagctct	ctctagcgac	ggccgcatct	tactgggtgt	4860
caatgtatat	cattttactg	ggggaccttg	tgcaagaactc	gtgggtgctgg	gcactgctgc	4920
tgctgcggca	gctggcaacc	tgacttgat	cgtcgcgac	ggaaatgaga	acaggggcat	4980
cttgagcccc	tgccgacggc	gccgacaggt	gcttctcgat	ctgcatcctg	ggatcaaagc	5040
catagtgaag	gacagtgatg	gacagccgac	ggcagttggg	attcgtgaat	tgctgccctc	5100
tggttatgtg	tgggagggct	aagcacaatt	cgagctcggc	acctttaaga	ccaatgactt	5160
acaaggcagc	tgtagatctt	agccactttt	taaaagaaaa	ggggggactg	gaagggctaa	5220
ttcactccca	acgaagacaa	gatctgcttt	ttgcttgtag	tgggtctctc	tggttagacc	5280

agatctgagc ctgggagctc tctggctaac tagggaaccc actgcttaag cctcaataaa	5340
gcttgcccttg agtgcttcaa gtagtgtgtg cccgtctgtt gtgtgactct ggtaactaga	5400
gatccctcag acccttttag tcagtgtgga aaatctctag cagtagtagt tcatgtcatc	5460
ttattattca gtatttataa cttgcaaaga aatgaatatac agagagttag aggaacttgt	5520
ttattgcagc ttataatggt tacaataaaa gcaatagcat cacaaatttc acaaataaag	5580
catttttttc actgcattct agttgtggtt tgtccaaact catcaatgta tcttatcatg	5640
tctggctcta gctatcccgc ccctaactcc gcccatcccg cccctaactc cgcccagttc	5700
cgcccatctc ccgccccatg gctgactaat tttttttatt tatgcagagg ccgaggccgc	5760
ctcggcctct gagctattcc agaagtagtg aggaggcttt tttggaggcc tagggacgta	5820
cccaattcgc cctatagtga gtcgtattac gcgcgctcac tggccgtcgt tttacaacgt	5880
cgtgactggg aaaaccctgg cgttacccaa cttaatcgcc ttgcagcaca tcccccttc	5940
gccagctggc gtaatagcga agaggcccg accgatcgcc cttcccaaca gttgcgcagc	6000
ctgaatggcg aatgggacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt	6060
acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc	6120
ccttcctttc tcgccacgtt cgccggcttt ccccgtaag ctctaaatcg ggggctccct	6180
ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat	6240
ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc	6300
acgttcttta atagtggact cttgttccaa actggaacaa cactcaaccc tatctcggtc	6360
tattcttttg atttataagg gattttgccg atttcggcct attgggttaa aaatgagctg	6420
atttaacaaa aatttaacgc gaattttaac aaaatattaa cgcttacaat ttaggtggca	6480
cttttcgggg aaatgtgcgc ggaacccta tttgtttatt tttctaaata cattcaaata	6540
tgtatccgct catgagacaa taaccctgat aaatgcttca ataattattga aaaaggaaga	6600
gtatgagtat tcaacatttc cgtgtcgcgc ttattccctt ttttgccgca ttttgccctc	6660
ctgtttttgc tcaccagaa acgctggtga aagtaaaaga tgctgaagat cagttgggtg	6720
cacgagtggg ttacatcgaa ctggatctca acagcggtaa gatccttgag agttttcgcc	6780
ccgaagaacg ttttccaatg atgagcactt ttaaagttct gctatgtggc gcggtattat	6840
cccgtattga cgccgggcaa gagcaactcg gtcgcgcgat acactattct cagaatgact	6900
tggttgagta ctcaccagtc acagaaaagc atcttacgga tggcatgaca gtaagagaat	6960
tatgcagtgc tgccataacc atgagtata acactgcggc caacttactt ctgacaacga	7020
tcggaggacc gaaggagcta accgcttttt tgcacaacat gggggatcat gtaactcgcc	7080
ttgatcggtg ggaaccggag ctgaatgaag ccataccaaa cgacgagcgt gacaccacga	7140

```

tgcctgtagc aatggcaaca acgttgcgca aactattaac tggcgaacta cttactctag 7200
cttcccggca acaattaata gactggatgg aggcggataa agttgcagga ccacttctgc 7260
gctcggccct tccggctggc tggtttattg ctgataaatc tggagccggt gagcgtgggt 7320
ctcgcggtat cattgcagca ctggggccag atggtaagcc ctcccgtatc gtagttatct 7380
acacgacggg gagtcaggca actatggatg aacgaaatag acagatcgct gagatagggtg 7440
cctcactgat taagcattgg taactgtcag accaagttta ctcatatata ctttagattg 7500
atttaaaact tcatttttaa tttaaaagga tctaggtgaa gatccttttt gataatctca 7560
tgacaaaaat cccttaacgt gagttttcgt tccactgagc gtcagacccc gtagaaaaga 7620
tcaaaggatc ttcttgagat cctttttttc tgcgcgtaat ctgctgcttg caaacaaaaa 7680
aaccaccgct accagcgggtg gtttgtttgc cggatcaaga gctaccaact ctttttccga 7740
aggtaactgg cttcagcaga gcgcagatac caaatactgt tcttctagt tagccgtagt 7800
taggccacca cttcaagaac tctgtagcac cgcctacata cctcgtcttg ctaatcctgt 7860
taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttggac tcaagacgat 7920
agttaccgga taaggcgag cggtcgggct gaacggggggg ttcgtgcaca cagcccagct 7980
tggagcgaac gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcca 8040
cgcttcccga agggagaaaag gcggacaggt atccggtaag cggcagggtc ggaacaggag 8100
agcgcacgag ggagcttcca gggggaaacg cctggtatct ttatagtcct gtcgggtttc 8160
gccacctctg acttgagcgt cgatttttgt gatgctcgtc agggggggcgg agcctatgga 8220
aaaacgccag caacgcggcc tttttacggt tcctggcctt ttgctggcct tttgctcaca 8280
tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagtgag 8340
ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg 8400
aagagcgccc aatacgcaaa ccgcctctcc ccgcgcgttg gccgattcat taatgcagct 8460
ggcacgacag gtttcccga c tggaaagcgg gcagtgagcg caacgcaatt aatgtgagtt 8520
agctcactca ttaggcaccc caggctttac actttatgct tccggctcgt atgttggtg 8580
gaattgtgag cggataacaa tttcacacag gaaacagcta tgaccatgat tacgccaagc 8640
gcgcaattaa ccctcactaa agggaacaaa agctggagct gcaagctt 8688

```

<210> 3

<211> 6964

<212> DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; pLenti6/V5-dTOPO™

&lt;400&gt; 3

aatgtagtct tatgcaatac tcttgtagtc ttgcaacatg gtaacgatga gtttagcaaca	60
tgccttacaa ggagagaaaa agcaccgtgc atgccgattg gtggaagtaa ggtggtacga	120
tcgtgcctta ttaggaaggc aacagacggg tctgacatgg attggacgaa ccactgaatt	180
gccgcattgc agagatattg tatttaagtg cctagctcga tacataaacg ggtctctctg	240
gtttagaccag atctgagcct gggagctctc tggctaacta gggaaccac tgcttaagcc	300
tcaataaagc ttgccttgag tgcttcaagt agtgtgtgcc cgtctgttgt gtgactctgg	360
taactagaga tccctcagac ccttttagtc agtgtggaaa atctctagca gtggcgcccg	420
aacagggact tgaaagcgaa agggaaacca gaggagctct ctgcagcgag gactcggctt	480
gctgaagcgc gcacggcaag aggcgagggg cggcgactgg tgagtacgcc aaaaattttg	540
actagcggag gctagaagga gagagatggg tgcgagagcg tcagtattaa gcgggggaga	600
attagatcgc gatgggaaaa aattcgggta aggccagggg gaaagaaaa atataaatta	660
aaacatatag tatgggcaag caggagagcta gaacgattcg cagttaatcc tggcctgtta	720
gaaacatcag aaggctgtag acaaatactg ggacagctac aaccatccct tcagacagga	780
tcagaagaac ttagatcatt atataataca gtagcaacct tctattgtgt gcatcaaagg	840
atagagataa aagacaccaa ggaagcttta gacaagatag aggaagagca aaacaaaagt	900
aagaccaccg cacagcaagc ggccgctgat cttcagacct ggaggaggag atatgaggga	960
caattggaga agtgaattat ataaatataa agtagtaaaa attgaaccat taggagtagc	1020
accaccaag gcaaagagaa gagtggtgca gagagaaaa agagcagtgg gaataggagc	1080
tttgttcctt gggttccttg gagcagcagg aagcactatg ggcgacgct caatgacgct	1140
gacggtacag gccagacaat tattgtctgg tatagtgcag cagcagaaca atttgctgag	1200
ggctattgag gcgcaacagc atctgttgca actcacagtc tggggcatca agcagctcca	1260
ggcaagaatc ctggctgtgg aaagatacct aaaggatcaa cagctcctgg ggatttgggg	1320
ttgctctgga aaactcattt gcaccactgc tgtgccttgg aatgctagtt ggagtaataa	1380
atctctggaa cagatttgga atcacacgac ctggatggag tgggacagag aaattaacaa	1440
ttacacaagc ttaatacact ccttaattga agaatcgcaa aaccagcaag aaaagaatga	1500
acaagaatta ttggaattag ataaatgggc aagtttgtgg aattggttta acataacaaa	1560

ttggctgtgg	tatataaaat	tattcataat	gatagtagga	ggcttggtag	gtttaagaat	1620
agtttttgct	gtactttcta	tagtgaatag	agttaggcag	ggatattcac	cattatcggt	1680
tcagaccac	ctcccaaccc	cgaggggacc	cgacaggccc	gaaggaatag	aagaagaagg	1740
tggagagaga	gacagagaca	gatccattcg	attagtgaac	ggatctcgac	ggtatcgata	1800
agcttgggag	ttccgcgtta	cataacttac	ggtaaattggc	ccgcctggct	gaccgcccaa	1860
cgacccccgc	ccattgacgt	caataatgac	gtatgttccc	atagtaacgc	caatagggac	1920
tttccattga	cgtcaatggg	tggagtattt	acggtaaact	gcccacttgg	cagtacatca	1980
agtgtatcat	atgccaagta	cgccccctat	tgacgtcaat	gacggtaaata	ggcccgctg	2040
gcattatgcc	cagtacatga	ccttatggga	ctttcctact	tggcagtaca	tctacgtatt	2100
agtcacgct	attaccatgg	tgatgcgggt	ttggcagtac	atcaatgggc	gtggatagcg	2160
gtttgactca	cggggatttc	caagtctcca	ccccattgac	gtcaatggga	gtttgttttg	2220
gcaccaaata	caacgggact	ttccaaaatg	tcgtaacaac	tccgccccat	tgacgcaaata	2280
gggcggtagg	cgtgtacggg	gggaggtcta	tataagcaga	gctcgtttag	tgaaccgtca	2340
gatcgccctg	agacgccatc	cacgctgttt	tgacctccat	agaagacacc	gactctagag	2400
gatccactag	tccagtgtgg	tggaaattgat	cccttcacca	agggctcgag	tctagagggc	2460
ccgcggttcg	aaggtaagcc	tatccctaac	cctctcctcg	gtctcgattc	tacgcgtacc	2520
ggtagtaata	gagtttgga	ttaattctgt	ggaatgtgtg	tcagttaggg	tgtggaaagt	2580
ccccaggctc	cccaggcagg	cagaagtatg	caaagcatgc	atctcaatta	gtcagcaacc	2640
aggtgtggaa	agtccccagg	ctccccagca	ggcagaagta	tgcaaagcat	gcattctcaat	2700
tagtcagcaa	ccatagtccc	gcccctaact	ccgcccattc	cgcccctaac	tccgcccagt	2760
tccgcccatt	ctccgcccc	tggctgacta	atTTTTTTT	tttatgcaga	ggccgaggcc	2820
gcctctgcct	ctgagctatt	ccagaagtag	tgaggaggct	tttttggagg	cctaggcttt	2880
tgcaaaaagc	tcccgggagc	ttgtatatcc	atTTTtcggat	ctgatcagca	cgtgttgaca	2940
attaatcatc	ggcatagtat	atcggcatag	tataatacga	caaggtagag	aactaaacca	3000
tggccaagcc	tttgtctcaa	gaagaatcca	ccctcattga	aagagcaacg	gtacaataca	3060
acagcatccc	catctctgaa	gactacagcg	tcgccagcgc	agctctctct	agcgacggcc	3120
gcattctcac	tgggtgtcaat	gtatatcatt	ttactggggg	accttgtgca	gaactcgtgg	3180
tgctgggcac	tgctgctgct	gcggcagctg	gcaacctgac	ttgtatcgtc	gcgatcgga	3240
atgagaacag	gggcatcttg	agccccctgc	gacgggtgcc	acagggtgct	ctcgatctgc	3300
atcctgggat	caaagccata	gtgaaggaca	gtgatggaca	gccgacggca	gttgggattc	3360
gtgaattgct	gccctctggt	tatgtgtggg	agggctaagc	acaattcgag	ctcggtacct	3420



ttaagaccaa	tgacttacaa	ggcagctgta	gatcttagcc	acttttttaa	agaaaagggg	3480
ggactggaag	ggctaattca	ctcccaacga	agacaagatc	tgctttttgc	ttgtactggg	3540
tctctctggt	tagaccagat	ctgagcctgg	gagctctctg	gctaactagg	gaaccactg	3600
cttaagcctc	aataaagctt	gccttgagtg	cttcaagtag	tgtgtgcccg	tctgttgtgt	3660
gactctggta	actagagatc	cctcagaccc	ttttagtcag	tgtggaaaat	ctctagcagt	3720
agtagttcat	gtcatcttat	tattcagtat	ttataacttg	caaagaaatg	aatatcagag	3780
agtgagagga	acttgtttat	tgcagcttat	aatgggttaca	aataaagcaa	tagcatcaca	3840
aatttcacaa	ataaagcatt	tttttcaactg	cattctagtt	gtggtttgtc	caaactcatc	3900
aatgtatctt	atcatgtctg	gctctagcta	tcccgccctt	aactccgccc	atcccgcccc	3960
taactccgcc	cagttccgcc	cattctccgc	cccatggctg	actaattttt	tttatttatg	4020
cagaggccga	ggccgcctcg	gcctctgagc	tattccagaa	gtagtgagga	ggcttttttg	4080
gaggcctagg	gacgtaccca	attcgcccta	tagtgagtcg	tattacgcgc	gctcactggc	4140
cgtcgtttta	caacgtcgtg	actgggaaaa	ccctggcggt	acccaactta	atcgccctgc	4200
agcacatccc	cctttcgcca	gctggcgtaa	tagcgaagag	gcccgcaccg	atcgcccttc	4260
ccaacagttg	cgcagcctga	atggcgaatg	ggacgcgccc	tgtagcggcg	cattaagcgc	4320
ggcgggtgtg	gtggttacgc	gcagcgtgac	cgctacactt	gccagcgcgc	tagcgcccgc	4380
tcctttcgct	ttcttccctt	cctttctcgc	cacgttcgcc	ggctttcccc	gtcaagctct	4440
aaatcggggg	ctcccttttag	ggttccgatt	tagtgcttta	cggcacctcg	accccaaaaa	4500
acttgattag	ggtgatgggt	cacgtagtgg	gccatcgccc	tgatagacgg	tttttcgccc	4560
tttgacgttg	gagtccacgt	tctttaatag	tggactcttg	ttccaaactg	gaacaacact	4620
caaccctatc	tcggtctatt	cttttgattt	ataagggatt	ttgccgattt	cggcctattg	4680
gttaaaaaat	gagctgattt	aacaaaaatt	taacgcgaat	tttaacaaaa	tattaacgct	4740
tacaatttag	gtggcacttt	tcggggaaat	gtgcgcggaa	cccctatttg	tttatttttc	4800
taaatacatt	caaatatgta	tccgctcatg	agacaataac	cctgataaat	gcttcaataa	4860
tattgaaaaa	ggaagagtat	gagtattcaa	catttccgtg	tcgcccttat	tccctttttt	4920
gcggcatttt	gccttcctgt	ttttgctcac	ccagaaacgc	tggtgaaagt	aaaagatgct	4980
gaagatcagt	tgggtgcacg	agtgggttac	atcgaactgg	atctcaacag	cggtaagatc	5040
cttgagagtt	ttcgccccga	agaacgtttt	ccaatgatga	gcacttttaa	agttctgcta	5100
tgtggcgcg	tattatcccc	tattgacgcc	gggcaagagc	aactcggtcg	ccgcatacac	5160
tattctcaga	atgacttggg	tgagtactca	ccagtcacag	aaaagcatct	tacggatggc	5220

atgacagtaa gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac	5280
ttacttctga caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg	5340
gatcatgtaa ctgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac	5400
gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc	5460
gaactactta ctctagcttc cgggcaacaa ttaatagact ggatggaggc ggataaagtt	5520
gcaggaccac ttctgcgctc ggcccttccg gctggctggc ttattgctga taaatctgga	5580
gccggtgagc gtgggtctcg cggtatcatt gcagcaactg ggccagatgg taagccctcc	5640
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag	5700
atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca	5760
tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc	5820
ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca	5880
gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc	5940
tgcttgcaaa caaaaaaacc accgctacca gcggtgggtt gtttgccgga tcaagagcta	6000
ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt	6060
ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc	6120
gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg	6180
ttggactcaa gacgatagtt accggataag gcgcagcggc cgggctgaac ggggggttcg	6240
tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag	6300
ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggatatcc ggtaagcggc	6360
agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat	6420
agtectgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg	6480
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc	6540
tggccttttg ctacatggtt ctttctgctg ttatcccctg attctgtgga taaccgtatt	6600
accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca	6660
gtgagcgagg aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttggccg	6720
attcattaat gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac	6780
gcaattaatg tgagttagct cactcattag gcaccccagg ctttacactt tatgcttccg	6840
gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac	6900
catgattacg ccaagcgcgc aattaaccct cactaaaggg aacaaaagct ggagctgcaa	6960
gctt	6964

&lt;210&gt; 4

&lt;211&gt; 8634

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; pLenti4/V5-DEST

&lt;400&gt; 4

aatgtagtct tatgcaatac tctttagtgc ttgcaacatg gtaacgatga gtttagcaaca	60
tgccttacaa ggagagaaaa agcaccgtgc atgccgattg gtggaagtaa ggtggtacga	120
tcgtagcctta ttaggaaggc aacagacggg tctgacatgg attggacgaa ccactgaatt	180
gccgcattgc agagatattg tatttaagtg cctagctcga tacataaacg ggtctctctg	240
gtttagaccag atctgagcct gggagctctc tggctaacta gggaaccac tgcttaagcc	300
tcaataaagc ttgccttgag tgcttcaagt agtgtgtgcc cgtctgttgt gtgactctgg	360
taactagaga tccctcagac ccttttagtc agtgtggaaa atctctagca gtggcgcccg	420
aacagggact tgaaagcgaa agggaaacca gaggagctct ctgcagcgag gactcggctt	480
gctgaagcgc gcacggcaag aggcgagggg cggcgactgg tgagtacgcc aaaaattttg	540
actagcggag gctagaagga gagagatggg tgcgagagcg tcagtattaa gcgggggaga	600
attagatcgc gatgggaaaa aattcgggta aggccagggg gaaagaaaa atataaatta	660
aaacatatag tatgggcaag caggagcta gaacgattcg cagttaatcc tggcctgtta	720
gaaacatcag aaggctgtag acaaatactg ggacagctac aaccatccct tcagacagga	780
tcagaagaac ttagatcatt atataatata gtagcaaccc tctattgtgt gcatcaaagg	840
atagagataa aagacaccaa ggaagcttta gacaagatag aggaagagca aaacaaaagt	900
aagaccaccg cacagcaagc ggccgctgat cttcagacct ggaggaggag atatgaggga	960
caattggaga agtgaattat ataaatataa agtagtaaaa attgaacct taggagtagc	1020
accaccaag gcaaagagaa gagtgggtgca gagagaaaa agagcagtgg gaataggagc	1080
tttgttcctt gggttcttgg gagcagcagg aagcactatg ggcgagcgt caatgacgct	1140
gacggtacag gccagacaat tattgtctgg tatagtgcag cagcagaaca atttgctgag	1200
ggctattgag gcgcaacagc atctgttgca actcacagtc tggggcatca agcagctcca	1260
ggcaagaatc ctggctgtgg aaagatacct aaaggatcaa cagctcctgg ggatttgggg	1320
ttgctctgga aaactcattt gcaccactgc tgtgccttgg aatgctagtt ggagtaataa	1380

atctctggaa	cagatttgga	atcacacgac	ctggatggag	tgggacagag	aaattaacaa	1440
ttacacaagc	ttaatacact	ccttaattga	agaatcgcaa	aaccagcaag	aaaagaatga	1500
acaagaatta	ttggaattag	ataaatgggc	aagtttgtgg	aattggttta	acataacaaa	1560
ttggctgtgg	tatataaaat	tattcataat	gatagtagga	ggcttggtag	gtttaagaat	1620
agtttttgct	gtactttcta	tagtgaatag	agttaggcag	ggatattcac	cattatcggt	1680
tcagaccac	ctcccaaccc	cgaggggacc	cgacaggccc	gaaggaatag	aagaagaagg	1740
tggagagaga	gacagagaca	gatccattcg	attagtgaac	ggatctcgac	ggtatcgata	1800
agcttgggag	ttcgcgtta	cataacttac	ggtaaattggc	ccgcctggct	gaccgcccaa	1860
cgacccccgc	ccattgacgt	caataatgac	gtatgttccc	atagtaacgc	caatagggac	1920
tttccattga	cgtcaatggg	tggagtattt	acggtaaact	gccacttgg	cagtacatca	1980
agtgtatcat	atgccaagta	cgccccctat	tgacgtcaat	gacggtaaata	ggcccgctg	2040
gcattatgcc	cagtacatga	ccttatggga	ctttcctact	tggcagtaca	tctacgtatt	2100
agtcacgct	attaccatgg	tgatgcgggt	ttggcagtac	atcaatgggc	gtggatagcg	2160
gtttgactca	cggggatttc	caagtctcca	ccccattgac	gtcaatggga	gtttgttttg	2220
gcacaaaaat	caacgggact	ttccaaaatg	tcgtaacaac	tccgccccat	tgacgcaaat	2280
gggcggtagg	cgtgtacggg	gggaggtcta	tataagcaga	gctcgtttag	tgaaccgtca	2340
gatcgccctg	agacgccatc	cacgctgttt	tgacctccat	agaagacacc	gactctagag	2400
gatccactag	tccagtgtgg	tggaaattctg	cagatatcaa	caagtttgta	caaaaaagct	2460
gaacgagaaa	cgtaaaatga	tataaatatc	aatatatata	attagatttt	gcataaaaaa	2520
cagactacat	aatactgtaa	aacacaacat	atccagtcac	tatggcggcc	gcattaggca	2580
ccccaggctt	tacactttat	gcttccggct	cgtataatgt	gtggattttg	agttaggatc	2640
cggcgagatt	ttcaggagct	aaggaagcta	aatggagaa	aaaaatcact	ggatatacca	2700
ccgttgatat	atcccaatgg	catcgtaaag	aacattttga	ggcatttcag	tcagttgctc	2760
aatgtaccta	taaccagacc	gttcagctgg	atattacggc	cttttttaaag	accgtaaaga	2820
aaaataagca	caagttttat	ccggccttta	ttcacattct	tgcccgcctg	atgaatgctc	2880
atccggaatt	ccgtatggca	atgaaagacg	gtgagctggg	gatatgggat	agtgttcacc	2940
cttgttacac	cgttttccat	gagcaaaactg	aaacgttttc	atcgctctgg	agtgaatacc	3000
acgacgattt	ccggcagttt	ctacacatat	attcgcaaga	tgtggcgtgt	tacggtgaaa	3060
acctggccta	tttccctaaa	gggtttattg	agaatatgtt	tttcgtctca	gccaatccct	3120
gggtgagttt	caccagtttt	gatttaaactg	tggccaatat	ggacaacttc	ttcgcccccg	3180
ttttcaccat	gggcaaatat	tatacgcaag	gcgacaaggt	gctgatgccg	ctggcgattc	3240

aggttcatca	tgccgtctgt	gatggcttcc	atgtcggcag	aatgcttaat	gaattacaac	3300
agtactgcga	tgagtggcag	ggcggggcgt	aaagatctgg	atccggctta	ctaaaagcca	3360
gataacagta	tgcgtatttg	cgcgctgatt	tttgcggtat	aagaatatat	actgatatgt	3420
ataccggaag	tatgtcaaaa	agaggtgtgc	tatgaagcag	cgtattacag	tgacagttga	3480
cagcgacagc	tatcagttgc	tcaaggcata	tatgatgtca	atatctccgg	tctggtaagc	3540
acaaccatgc	agaatgaagc	ccgtcgtctg	cgtgccgaac	gctggaaagc	ggaaaatcag	3600
gaagggatgg	ctgaggtcgc	ccggtttatt	gaaatgaacg	gctcttttgc	tgacgagaac	3660
agggactggg	gaaatgcagt	ttaaggttta	cacctataaa	agagagagcc	gttatcgtct	3720
gtttgtggat	gtacagagtg	atattattga	cacgcccggg	cgacggatgg	tgatccccct	3780
ggccagtgca	cgtctgctgt	cagataaagt	ctcccgtgaa	ctttaccggg	tggtgcatat	3840
cggggatgaa	agctggcgca	tgatgaccac	cgatatggcc	agtgtgccgg	tctccgttat	3900
cggggaagaa	gtggctgata	tcagccaccg	cgaaaatgac	atcaaaaacg	ccattaacct	3960
gatgttctgg	ggaatataaa	tgtcaggctc	cgttatacac	agccagtctg	caggtcgacc	4020
atagtgactg	gatatgttgt	gttttacagt	attatgtagt	ctgtttttta	tgcaaaatct	4080
aatttaatat	attgatattt	atatcatttt	acgtttctcg	ttcagctttc	ttgtacaaag	4140
tggttgatat	ccagcacagt	ggcggccgct	cgagtctaga	gggcccgcgg	ttcgaaggta	4200
agcctatccc	taaccctctc	ctcggctctg	attctacgcg	taccggttag	taatgagttt	4260
ggaattaatt	ctgtggaatg	tgtgtcagtt	aggggtgtgga	aagtccccag	gctccccagg	4320
caggcagaag	tatgcaaagc	atgcatctca	attagtcagc	aaccagggtg	ggaaagtccc	4380
caggctcccc	agcaggcaga	agtatgcaaa	gcatgcatct	caattagtca	gcaaccatag	4440
tcccggcccct	aactccgccc	atcccgcgcc	taactccgcc	cagttccgcc	cattctccgc	4500
cccatggctg	actaattttt	tttatttatg	cagaggccga	ggccgcctct	gcctctgagc	4560
tattccagaa	gtagtgagga	ggcttttttg	gaggcctagg	cttttgcaaa	aagctcccc	4620
tggtgacaat	taatcatcgg	catagtatat	cggcatagta	taatacgaca	aggtgaggaa	4680
ctaaaccatg	gccaagttga	ccagtgccgt	tccgggtgctc	accgcgcgcg	acgtcgccgg	4740
agcggtcgag	ttctggaccg	accggctcgg	gttctcccgg	gacttcgtgg	aggacgactt	4800
cgccgggtgtg	gtccgggacg	acgtgaccct	gttcatcagc	gcgggtccagg	accagggtgg	4860
gccggacaac	accctggcct	gggtgtgggt	gcgcggcctg	gacgagctgt	acgccgagtg	4920
gtcggaggtc	gtgtccacga	acttccggga	cgcctccggg	ccggccatga	ccgagatcgg	4980
cgagcagccg	tgggggcggg	agttcgccct	gcgcgaccgc	gccggcaact	gcgtgcactt	5040

cgtggccgag gagcaggact gacacgtgct acgagattta aatgggtacct ttaagaccaa	5100
tgacttacaa ggcagctgta gatcttagcc acttttttaa agaaaagggg ggactggaag	5160
ggctaattca ctcccaacga agacaagatc tgctttttgc ttgtactggg tctctctggt	5220
tagaccagat ctgagcctgg gagctctctg gctaactagg gaaccactg cttagcctc	5280
aataaagctt gccttgagtg cttcaagtag tgtgtgcccg tctgttgtgt gactctggta	5340
actagagatc cctcagaccc ttttagtcag tgtggaaaat ctctagcagt agtagttcat	5400
gtcatcttat tattcagtat ttataacttg caaagaaatg aatatcagag agtgagagga	5460
acttgtttat tgcagcttat aatgggttaca aataaagcaa tagcatcaca aatttcacaa	5520
ataaagcatt tttttcactg cattctagtt gtgggttgtc caaactcatc aatgtatctt	5580
atcatgtctg gctctagcta tcccgcccct aactccgcc atcccgcccc taactccgcc	5640
cagttccgcc cattctccgc cccatggctg actaattttt tttatttatg cagaggccga	5700
ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggcttttttg gaggcctagg	5760
gacgtacca attcgcccta tagtgagtcg tattacgcgc gctcactggc cgtcgtttta	5820
caacgtcgtg actgggaaaa ccctggcggt acccaactta atcgccctgc agcacatccc	5880
cctttcgcca gctggcgtaa tagcgaagag gcccgcaccc atcgcccttc ccaacagttg	5940
cgcagcctga atggcgaaat ggacgcgccc tgtagcggcg cattaagcgc ggcggtgtg	6000
gtggttacgc gcagcgtgac cgctacactt gccagcgccc tagcgccgcg tcctttcgct	6060
ttcttccctt cctttctcgc cacgttcgcc ggctttcccc gtcaagctct aaatcggggg	6120
ctccctttag ggttccgatt tagtgcttta cggcacctcg accccaaaaa acttgattag	6180
ggtgatgggt cacgtagtgg gccatcgccc tgatagacgg tttttcgccc ttgacgttg	6240
gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact caaccctatc	6300
tcggtctatt cttttgatth ataagggatt ttgccgattt cggcctattg gttaaaaaat	6360
gagctgattt aacaaaaatt taacgcgaat tttaacaaaa tattaacgct tacaatttag	6420
gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt	6480
caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa	6540
ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt	6600
gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt	6660
tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaaagatc cttgagagtt	6720
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg	6780
tattatcccc tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga	6840
atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa	6900

gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga	6960
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa	7020
ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca	7080
ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta	7140
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac	7200
ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc	7260
gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag	7320
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga	7380
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt	7440
agattgattht aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata	7500
atctcatgac caaaatccct taacgtgagt tttcggtcca ctgagcgtca gaccccgtag	7560
aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa	7620
caaaaaaacc accgctacca gcggtgggtt gtttgccgga tcaagagcta ccaactcttt	7680
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc	7740
cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa	7800
tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa	7860
gacgatagtt accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc	7920
ccagcttggg gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa	7980
gcgccacgct tcccgaaggg agaaaggcgg acaggatatcc ggtaagcggc agggtcggaa	8040
caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg	8100
ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc	8160
tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg	8220
ctcacatggt ctttcctgcg ttatcccttg attctgtgga taaccgtatt accgcctttg	8280
agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcagatca gtgagcgagg	8340
aagcggaga gcgccaata cgcaaaccgc ctctccccgc gcgttggccg attcattaat	8400
gcagctggca cgacagggtt cccgactgga aagcgggcag tgagcgcaac gcaattaatg	8460
tgagttagct cactcattag gcaccccagg ctttacactt tatgcttccg gctcgtatgt	8520
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg	8580
ccaagcgcgc aattaaccct cactaaaggg aacaaaagct ggagctgcaa gctt	8634

&lt;210&gt; 5

&lt;211&gt; 9320

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; pLenti6/Ubc/V5-DEST

&lt;400&gt; 5

aatgtagtct tatgcaatac tcttgtagtc ttgcaacatg gtaacgatga gttagcaaca	60
tgccttaciaa ggagagaaaa agcaccgtgc atgccgattg gtggaagtaa ggtggtacga	120
tcgtgcctta ttaggaaggc aacagacggg tctgacatgg attggacgaa ccaactgaatt	180
gccgcattgc agagatatgg tatttaagtg cctagctcga tacataaacg ggtctctctg	240
gttagaccag atctgagcct gggagctctc tggctaacta gggaaccacac tgcttaagcc	300
tcaataaagc ttgccttgag tgcttcaagt agtgtgtgcc cgtctgttgt gtgactctgg	360
taactagaga tccctcagac ccttttagtc agtgtggaaa atctctagca gtggcgcccg	420
aacagggact tgaaagcgaa agggaaacca gaggagctct ctgcacgcag gactcggctt	480
gctgaagcgc gcacggcaag aggcgagggg cggcgactgg tgagtacgcc aaaaattttg	540
actagcggag gctagaagga gagagatggg tgcgagagcg tcagtattaa gcgggggaga	600
attagatcgc gatgggaaaa aattcgggta aggccagggg gaaagaaaaa atataaatta	660
aaacatatag tatgggcaag caggagctga gaacgattcg cagttaatcc tggcctgtta	720
gaaacatcag aaggctgtag acaaatactg ggacagctac aaccatccct tcagacagga	780
tcagaagaac ttagatcatt atataatata gtagcaaccc tctatttgtgt gcatcaaagg	840
atagagataa aagacaccaa ggaagcttta gacaagatag aggaagagca aaacaaaagt	900
aagaccaccg cacagcaagc ggccgctgat cttcagacct ggaggaggag atatgaggga	960
caattggaga agtgaattat ataaatataa agtagtaaaa attgaacat taggagtagc	1020
accaccaag gcaaagagaa gagtggtgca gagagaaaaa agagcagtgg gaataggagc	1080
ttgtttcctt gggttcttgg gagcagcagg aagcactatg ggcgacgcgt caatgacgct	1140
gacggtacag gccagacaat tattgtctgg tatagtgcag cagcagaaca atttgctgag	1200
ggctattgag gcgcaacagc atctgttgca actcacagtc tggggcatca agcagctcca	1260
ggcaagaatc ctggctgtgg aaagatacct aaaggatcaa cagctcctgg ggatttgagg	1320
ttgctctgga aaactcattt gcaccactgc tgtgccttgg aatgctagtt ggagtaataa	1380



atctctggaa	cagatttgga	atcacacgac	ctggatggag	tgggacagag	aaattaacaa	1440
ttacacaagc	ttaatacact	ccttaattga	agaatcgcaa	aaccagcaag	aaaagaatga	1500
acaagaatta	ttggaattag	ataaatgggc	aagtttgtgg	aattggttta	acataacaaa	1560
ttggctgtgg	tatataaaat	tattcataat	gatagtagga	ggcttggtag	gtttaagaat	1620
agtttttgct	gtactttcta	tagtgaatag	agttaggcag	ggatattcac	cattatcggt	1680
tcagaccac	ctcccaaccc	cgaggggacc	cgacaggccc	gaaggaatag	aagaagaagg	1740
tggagagaga	gacagagaca	gatccattcg	attagtgaac	ggatctcgac	ggtatcggat	1800
ctggcctccg	cgccggggtt	tggcgccctc	cgccggcgcc	cccctcctca	cggcgagcgc	1860
tgccacgtca	gacgaagggc	gcaggagcgt	cctgatcctt	ccgcccggac	gctcaggaca	1920
gcggcccgt	gctcataaga	ctcggcctta	gaaccccagt	atcagcagaa	ggacatttta	1980
ggacgggact	tgggtgactc	tagggcactg	gttttctttc	cagagagcgg	aacaggcgag	2040
gaaaagtagt	cccttctcgg	cgattctgcg	gagggatctc	cgtggggcgg	tgaacgccga	2100
tgattatata	aggacgcgcc	gggtgtggca	cagctagttc	cgtcgcagcc	gggatttggg	2160
tcgcggttct	tgtttgtgga	tcgctgtgat	cgtcacttgg	tgagtagcgg	gctgctgggc	2220
tggccggggc	tttcgtggcc	gccgggcccgc	tcggtgggac	ggaagcgtgt	ggagagaccg	2280
ccaagggctg	tagtctgggt	ccgcgagcaa	ggttgccctg	aactgggggt	tggggggagc	2340
gcagcaaaat	ggcggctgtt	cccaggtctt	gaatggaaga	cgcttgtgag	gcgggctgtg	2400
aggtcgttga	aacaaggtgg	ggggcatggg	gggcggcaag	aacccaaggt	cttgaggcct	2460
tcgctaattgc	gggaaagctc	ttattcgggt	gagatgggct	ggggcaccat	ctggggaccc	2520
tgacgtgaag	tttgtcactg	actggagaac	tcggtttgtc	gtctgttgcg	ggggcggcag	2580
ttatgcgggtg	ccgttgggca	gtgcacccgt	acctttggga	gcgcgcgccc	tcgtcgtgtc	2640
gtgacgtcac	ccgttctgtt	ggcttataat	gcagggtggg	gccacctgcc	ggtaggtgtg	2700
cggtaggctt	ttctccgtcg	caggacgcag	ggttcgggcc	tagggtaggc	tctcctgaat	2760
cgacaggcgc	cggacctctg	gtgaggggag	ggataagtga	ggcgtcagtt	tctttggtcg	2820
gttttatgta	cctatcttct	taagtagctg	aagctccggg	tttgaactat	gcgctcgggg	2880
ttggcgagtg	tgttttgtga	agtttttttag	gcaccttttg	aaatgtaatc	atttgggtca	2940
atatgtaatt	ttcagtgtta	gactagtaaa	ttgtccgcta	aattctggcc	gtttttggct	3000
tttttggttag	acgaagcttg	gtaccgagct	cggatccact	agtccagtgt	ggtggaattc	3060
tgcagatatc	aacaagtttg	tacaaaaaag	ctgaacgaga	aacgtaaaat	gatataaata	3120
tcaatatatt	aaattagatt	ttgcataaaa	aacagactac	ataatactgt	aaaacacaac	3180
atatccagtc	actatggcgg	ccgcattagg	caccccaggc	tttacacttt	atgcttccgg	3240

ctctgtataat gtgtggattt tgagtttagga tccggcgaga ttttcaggag ctaaggaagc 3300  
 taaaatggag aaaaaaatca ctggatatac caccgttgat atatcccaat ggcatcgtaa 3360  
 agaacatttt gaggcatttc agtcagttgc tcaatgtacc tataaccaga ccgttcagct , 3420  
 ggatattacg gccttttttaa agaccgtaaa gaaaaataag cacaagtttt atccggcctt 3480  
 tattcacatt cttgcccgcc tgatgaatgc tcatccggaa ttccgtatgg caatgaaaga 3540  
 cggtgagctg gtgatatggg atagtgttca cccttgttac accgttttcc atgagcaaac 3600  
 tgaaacgttt tcatcgctct ggagtgaata ccacgacgat ttccggcagt ttctacacat 3660  
 atattcgcaa gatgtggcgt gttacgggtga aaacctggcc tatttcccta aagggtttat 3720  
 tgagaatatg tttttcgtct cagccaatcc ctgggtgagt ttcaccagtt ttgatttaaa 3780  
 cgtggccaat atggacaact tcttcgcccc cgttttcacc atgggcaaatt attatacgca 3840  
 aggcgacaag gtgctgatgc cgctggcgat tcaggttcat catgccgtct gtgatggctt 3900  
 ccatgtcggc agaatgctta atgaattaca acagtactgc gatgagtggc agggcggggc 3960  
 gtaaagatct ggatccggct tactaaaagc cagataacag tatgcgtatt tgcgcgctga 4020  
 tttttgcggt ataagaatat atactgatat gtatacccgga agtatgtcaa aaagaggtgt 4080  
 gctatgaagc agcgtattac agtgacagtt gacagcgaca gctatcagtt gctcaaggca 4140  
 tatatgatgt caatatctcc ggtctggtaa gcacaaccat gcagaatgaa gcccgtcgtc 4200  
 tgcgtgccga acgctggaaa gcggaaaatc aggaagggat ggctgaggtc gcccggttta 4260  
 ttgaaatgaa cggctctttt gctgacgaga acagggactg gtgaaatgca gtttaagggt 4320  
 tacacctata aaagagagag ccgttatcgt ctgtttgtgg atgtacagag tgatattatt 4380  
 gacacgcccg ggcgacggat ggtgatcccc ctggccagtg cacgtctgct gtcagataaa 4440  
 gtctcccgtg aactttaccc ggtgggtgcat atcgggggatg aaagctggcg catgatgacc 4500  
 accgatatgg ccagtgtgcc ggtctccgtt atcgggggaag aagtggctga tctcagccac 4560  
 cgcgaaaatg acatcaaaaa cgccattaac ctgatgttct ggggaatata aatgtcaggc 4620  
 tccgttatac acagccagtc tgcaggtcga ccatagtgac tggatatgtt gtgttttaca 4680  
 gtattatgta gtctgttttt tatgcaaaat ctaatttaat atattgatat ttatatcatt 4740  
 ttacgtttct cgttcagctt tcttgtaaa agtggttgat atccagcaca gtggcggccg 4800  
 ctcgagtcta gagggcccgc ggttcgaagg taagcctatc cctaaccctc tcctcggctc 4860  
 cgattctacg cgtaccggtt agtaatgagt ttggaattaa ttctgtggaa tgtgtgtcag 4920  
 ttaggggtgtg gaaagtcccc aggcctccca ggcaggcaga agtatgcaa gcattgcatct 4980  
 caattagtca gcaaccaggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca 5040

aagcatgcat	ctcaattagt	cagcaaccat	agtccccgcc	ctaactccgc	ccatccccgcc	5100
cctaactccg	cccagttccg	cccattctcc	gccccatggc	tgactaattt	tttttattta	5160
tgcagaggcc	gaggccgcct	ctgcctctga	gctattccag	aagtagtgag	gaggcttttt	5220
tggaggccta	ggcttttgca	aaaagctccc	gggagcttgt	atatccattt	tcggatctga	5280
tcagcacgtg	ttgacaatta	atcatcggca	tagtatatcg	gcatagtata	atacgacaag	5340
gtgaggaact	aaaccatggc	caagcctttg	tctcaagaag	aatccaccct	cattgaaaga	5400
gcaacggcta	caatcaacag	catccccatc	tctgaagact	acagcgtcgc	cagcgcagct	5460
ctctctagcg	acggccgcag	cttcactggg	gtcaatgtat	atcattttac	tgggggacct	5520
tgtgcagaac	tcgtgggtgt	gggcactgct	gctgctgcgg	cagctggcaa	cctgacttgt	5580
atcgtcgcga	tcggaaatga	gaacaggggc	atcttgagcc	cctgcggacg	gtgccgacag	5640
gtgcttctcg	atctgcatcc	tgggatcaaa	gccatagtga	aggacagtga	tggacagccg	5700
acggcagttg	ggattcgtga	attgctgccc	tctggttatg	tgtgggaggg	ctaagcacia	5760
ttcgagctcg	gtaccttta	gaccaatgac	ttacaaggca	gctgtagatc	ttagccactt	5820
tttaaaagaa	aaggggggac	tggaagggct	aattcactcc	caacgaagac	aagatctgct	5880
ttttgcttgt	actgggtctc	tctggttaga	ccagatctga	gcctgggagc	tctctggcta	5940
actagggaac	ccactgctta	agcctcaata	aagcttgcc	tgagtgcctc	aagtagtgtg	6000
tgcccgtctg	ttgtgtgact	ctggtaacta	gagatccctc	agaccctttt	agtcagtggtg	6060
gaaaatctct	agcagtagta	gttcatgtca	tcttattatt	cagtatttat	aacttgcaaa	6120
gaaatgaata	tcagagagtg	agaggaactt	gtttattgca	gcttataatg	gttacaaata	6180
aagcaatagc	atcaciaaatt	tcaciaataa	agcatttttt	tcactgcatt	ctagttgtgg	6240
tttgtccaaa	ctcatcaatg	tatcttatca	tgtctggctc	tagctatccc	gcccctaact	6300
ccgcccattc	cgcccctaac	tccgcccagt	tccgcccatt	ctccgcccc	tggctgacta	6360
atTTTTTTT	tttatgcaga	ggccgaggcc	gcctcggcct	ctgagctatt	ccagaagtag	6420
tgaggaggct	tttttgagg	cctagggacg	taccaaatc	gccctatagt	gagtcgtatt	6480
acgcgcgctc	actggccgct	gttttacaac	gtcgtgactg	ggaaaaccct	ggcgttaccc	6540
aacttaatcg	ccttgacgca	catccccctt	tcgccagctg	gcgtaatagc	gaagaggccc	6600
gcaccgatcg	cccttcccaa	cagttgcgca	gcctgaatgg	cgaatgggac	gcgccctgta	6660
gcggcgcatt	aagcgcggcg	gggtgtgggtg	ttacgcgcag	cgtgaccgct	acacttgcca	6720
gcgccctagc	gcccgtcct	ttcgctttct	tcccttcctt	tctcgccacg	ttcgccggct	6780
ttccccgtca	agctctaaat	cgggggctcc	ctttagggtt	ccgatttagt	gctttacggc	6840
acctcgaccc	caaaaaactt	gattagggtg	atggttcacg	tagtgggcca	tcgccctgat	6900

agacggtttt	tcgccctttg	acgttggagt	ccacgttctt	taatagtga	ctcttgttcc	6960
aaactggaac	aacactcaac	cctatctcgg	tctattcttt	tgatttataa	gggattttgc	7020
cgatttcggc	ctattggtta	aaaaatgagc	tgatttaaca	aaaatttaac	gcgaatttta	7080
acaaaatatt	aacgcttaca	atttaggtgg	cacttttcgg	ggaaatgtgc	gcggaacccc	7140
tatttgttta	tttttctaaa	tacattcaaa	tatgtatccg	ctcatgagac	aataaccctg	7200
ataaatgctt	caataatatt	gaaaaaggaa	gagtatgagt	attcaacatt	tccgtgtcgc	7260
ccttattccc	ttttttgcgg	cattttgcct	tcctgttttt	gctcaccag	aaacgctggt	7320
gaaagtaaaa	gatgctgaag	atcagttggg	tgcacgagtg	ggttacatcg	aactggatct	7380
caacagcgg	aagatccttg	agagttttcg	ccccgaagaa	cgttttccaa	tgatgagcac	7440
ttttaagtt	ctgctatgtg	gcgcggtatt	atcccgtatt	gacgccgggc	aagagcaact	7500
cggtcgccgc	atacactatt	ctcagaatga	cttggttgag	tactcaccag	tcacagaaaa	7560
gcatcttacg	gatggcatga	cagtaagaga	attatgcagt	gctgccataa	ccatgagtga	7620
taacactgcg	gccaacttac	ttctgacaac	gatcggagga	ccgaaggagc	taaccgcttt	7680
tttgcacaa	atgggggatc	atgtaactcg	ccttgatcgt	tggaaccgg	agctgaatga	7740
agccatacca	aacgacgagc	gtgacaccac	gatgcctgta	gcaatggcaa	caacgttgcg	7800
caaactatta	actggcgaac	tacttactct	agcttcccgg	caacaattaa	tagactggat	7860
ggaggcggat	aaagttgcag	gaccacttct	gcgctcggcc	cttccggctg	gctggtttat	7920
tgctgataaa	tctggagccg	gtgagcgtgg	gtctcgcggt	atcattgcag	cactggggcc	7980
agatggtaag	ccctcccgtg	tcgtagtatt	ctacacgacg	gggagtcagg	caactatgga	8040
tgaacgaaat	agacagatcg	ctgagatagg	tgccctactg	attaagcatt	ggtaactgtc	8100
agaccaagtt	tactcatata	tacttttagat	tgatttaaaa	cttcattttt	aatttaaaag	8160
gatctaggtg	aagatccttt	ttgataatct	catgaccaaa	atcccttaac	gtgagttttc	8220
gttccactga	gcgtcagacc	ccgtagaaaa	gatcaaagga	tcttcttgag	atcctttttt	8280
tctgcgcgta	atctgctgct	tgcaaacaaa	aaaaccaccg	ctaccagcgg	tggtttgttt	8340
gccggatcaa	gagctaccaa	ctcttttttc	gaaggtaact	ggcttcagca	gagcgcagat	8400
accaaatact	gttcttctag	tgtagccgta	gttaggccac	cacttcaaga	actctgtagc	8460
accgcctaca	tacctcgctc	tgctaatact	gttaccagtg	gctgctgcca	gtggcgataa	8520
gtcgtgtctt	accgggttgg	actcaagacg	atagttaccg	gataaggcgc	agcggtcggg	8580
ctgaacgggg	ggttcgtgca	cacagcccag	cttgagagca	acgacctaca	ccgaactgag	8640
atacctacag	cgtgagctat	gagaaagcgc	cacgcttccc	gaaggagaaa	aggcggacag	8700

gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa	8760
cgcttggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt	8820
gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg	8880
gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc	8940
tgtggataac cgtattaccg cctttgagtg agctgatacc gctcgccgca gccgaacgac	9000
cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct	9060
ccccgcgcgt tggcogattc attaatgcag ctggcacgac aggtttcccg actggaaagc	9120
gggcagtgag cgcaacgcaa ttaatgtgag ttagctcact cattagggac cccaggcttt	9180
acactttatg cttccggctc gtatgttggtg tggaattgtg agcggataac aatttcacac	9240
aggaaacagc tatgaccatg attacgcaa gcgcgcaatt aaccctcact aaaggaaca	9300
aaagctggag ctgcaagctt	9320

&lt;210&gt; 6

&lt;211&gt; 8889

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; pLP1

&lt;400&gt; 6

ttggccatt gcatacgttg tatccatata ataatatgta catttatatt ggctcatgtc	60
caacattacc gccatgttga cattgattat tgactagtta ttaatagtaa tcaattacgg	120
ggtcattagt tcatagccca tatatggagt tccgcgttac ataacttacg gtaaattggc	180
cgcttggtg accgcccaac gacccccgcc cattgacgtc aataatgacg tatgttcca	240
tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg	300
cccacttggc agtacatcaa gtgtatcata tgccaagtac gcccctatt gacgtcaatg	360
acggtaaattg gccgcctgg cattatgccc agtacatgac cttatgggac tttcctactt	420
ggcagtacat ctacgtatta gtcacgcta ttaccatggg gatgcgggtt tggcagtaca	480
tcaatgggcg tggatagcgg tttgactcac ggggatttcc aagtctccac cccattgacg	540
tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt cgtaacaact	600
ccgccccatt gacgcaaattg ggcggtaggc gtgtacggtg ggaggtctat ataagcagag	660
ctcgttttagt gaaccgtcag atcgccctgga gacgccatcc acgctgtttt gacctccata	720

gaagacaccg ggaccgatcc agcctcccct cgaagcttac atgtggtacc gagctcggat	780
cctgagaact tcaggggtgag tctatgggac ccttgatggt ttctttcccc ttcttttcta	840
tggttaagtt catgtcatag gaaggggaga agtaacaggg tacacatatt gaccaaataca	900
gggtaatttt gcatttgtaa ttttaaaaaa tgctttcttc ttttaatatata cttttttggt	960
tatcttattt ctaatacttt ccctaattct tttctttcag ggcaataatg atacaatgta	1020
tcatgcctct ttgcaccatt ctaaagaata acagtataa tttctgggtt aaggcaatag	1080
caatatttct gcatataaat atttctgcat ataaattgta actgatgtaa gaggtttcat	1140
attgctaata gcagctacaa tccagctacc attctgcttt tatttttatgg ttgggataag	1200
gctggattat tctgagtcca agctaggccc ttttgctaata catgttcata cctcttatct	1260
tcctcccaca gctcctgggc aacgtgctgg tctgtgtgct ggcccatcac tttggcaaag	1320
cacgtgagat ctgaattcga gatctgccgc cgccatgggt gcgagagcgt cagtattaag	1380
cggggggagaa ttagatcgat gggaaaaaat tcggttaagg ccaggggggaa agaaaaata	1440
taaattaaaa catatagtat gggcaagcag ggagctagaa cgattcgcag ttaatcctgg	1500
cctgttagaa acatcagaag gctgtagaca aatactggga cagctacaac catcccttca	1560
gacaggatca gaagaactta gatcattata taatacagta gcaaccctct attgtgtgca	1620
tcaaaggata gagataaaaag acaccaagga agcttttagac aagatagagg aagagcaaaa	1680
caaaagtaag aaaaagcac agcaagcagc agctgacaca ggacacagca atcaggtcag	1740
ccaaaattac cctatagtgc agaacatcca ggggcaaata gtacatcagg ccatatcacc	1800
tagaacttta aatgcatggg taaaagtagt agaagagaag gctttcagcc cagaagtgat	1860
acccatgttt tcagcattat cagaaggagc caccaccaca gatttaaaca ccatgctaaa	1920
cacagtgggg ggacatcaag cagccatgca aatgttaaaa gagaccatca atgaggaagc	1980
tgcagaatgg gatagagtgc atccagtgc tgcagggcct attgcaccag gccagatgag	2040
agaaccaagg ggaagtgaca tagcaggaac tactagtacc cttcaggaac aaataggatg	2100
gatgacacat aatccaccta tcccagtagg agaaatctat aaaagatgga taatcctggg	2160
attaaataaa atagtaagaa tgtatagccc taccagcatt ctggacataa gacaaggacc	2220
aaaggaaccc tttagagact atgtagaccg attctataaa actctaagag ccgagcaagc	2280
ttcacaagag gtaaaaaatt ggatgacaga aaccttggtg gtccaaaatg cgaaccaga	2340
ttgtaagact attttaaaag cattgggacc aggagcgaca ctagaagaaa tgatgacagc	2400
atgtcagggg gtgggggggac ccggccataa agcaagagtt ttggctgaag caatgagcca	2460
agtaacaaat ccagctacca taatgatata gaaaggcaat tttaggaacc aaagaaagac	2520

tgттаagtgt ttcaattgtg gcaaagaagg gcacatagcc aaaaattgca gggcccctag	2580
gaaaaagggc tgttggaat gtggaaagga aggacaccaa atgaaagatt gtactgagag	2640
acaggctaат tttttaggga agatctggcc ttcccacaag ggaaggccag ggaattttct	2700
tcagagcaga ccagagccaa cagccccacc agaagagagc ttcaggtttg gggaagagac	2760
aacaactccc tctcagaagc aggagccgat agacaaggaa ctgtatcctt tagcttccct	2820
cagatcactc tttggcagcg acccctcgtc acaataaaga taggggggca attaaaggaa	2880
gctctattag atacaggagc agatgataca gtattagaag aaatgaattt gccaggaaga	2940
tggaaccaa aaatgatagg gggaattgga ggttttatca aagtaagaca gtatgatcag	3000
atactcatag aaatctgcg acataaagct ataggtacag tattagtagg acctacacct	3060
gtcaacataa ttggaagaaa tctgttgact cagattggct gcactttaaa ttttcccatt	3120
agtcctattg agactgtacc agtaaaatta aagccaggaa tggatggccc aaaagttaaa	3180
caatggccat tgacagaaga aaaaataaaa gcattagtag aaatttgtac agaaatggaa	3240
aaggaaggaa aaatttcaaa aattgggcct gaaaatccat acaatactcc agtatttgcc	3300
ataaagaaaa aagacagtac taaatggaga aaattagtag atttcagaga acttaataag	3360
agaactcaag atttctggga agttcaatta ggaataccac atcctgcagg gttaaaacag	3420
aaaaaatcag taacagtact ggatgtgggc gatgcatatt tttcagttcc cttagataaa	3480
gacttcagga agtatactgc atttaccata cctagtataa acaatgagac accagggatt	3540
agatatcagt acaatgtgct tccacaggga tggaaaggat caccagcaat attccagtgt	3600
agcatgacaa aaatcttaga gccttttaga aaacaaaatc cagacatagt catctatcaa	3660
tacatggatg atttgtatgt aggatctgac ttagaaatag ggcagcatag aacaaaaata	3720
gaggaactga gacaacatct gttgaggtgg ggatttacca caccagacaa aaaacatcag	3780
aaagaacctc cattcctttg gatgggttat gaactccatc ctgataaatg gacagtacag	3840
cctatagtgc tgccagaaaa ggacagctgg actgtcaatg acatacagaa attagtggga	3900
aaattgaatt gggcaagtca gatttatgca gggattaaag taaggcaatt atgtaaactt	3960
cttaggggaa ccaaagcact aacagaagta gtaccactaa cagaagaagc agagctagaa	4020
ctggcagaaa acagggagat tctaaaagaa ccggtacatg gagtgtatta tgacccatca	4080
aaagacttaa tagcagaaat acagaagcag gggcaaggcc aatggacata tcaaatttat	4140
caagagccat ttaaaaatct gaaaacagga aagtatgcaa gaatgaaggg tgcccacact	4200
aatgatgtga aacaattaac agaggcagta caaaaaatag ccacagaaag catagtaata	4260
tggggaaaga ctctaaatt taaattaccc atacaaaagg aaacatggga agcatgggtg	4320
acagagtatt ggcaagccac ctggattcct gagtgggagt ttgtcaatac cctccctta	4380

gtgaagttat ggtaccagtt agagaaagaa cccataatag gagcagaaac tttctatgta	4440
gatggggcag ccaataggga aactaaatta ggaaaagcag gatatgtaac tgacagagga	4500
agacaaaaag ttgtccccct aacggacaca acaaatcaga agactgagtt acaagcaatt	4560
catctagctt tgcaggattc gggattagaa gtaaacaatag tgacagactc acaatatgca	4620
ttgggaatca ttcaagcaca accagataag agtgaatcag agttagtcag tcaaataata	4680
gagcagttaa taaaaaagga aaaagtctac ctggcatggg taccagcaca caaaggaatt	4740
ggaggaaatg aacaagtaga taaattggtc agtgctggaa tcaggaaagt actattttta	4800
gatggaatag ataaggccca agaagaacat gagaaatatc acagtaattg gagagcaatg	4860
gctagtgatt ttaacctacc acctgtagta gcaaaagaaa tagtagccag ctgtgataaa	4920
tgtcagctaa aaggggaagc catgcatgga caagtagact gtagcccagg aatatggcag	4980
ctagattgta cacatttaga aggaaaagtt atcttggtag cagttcatgt agccagtgga	5040
tatatagaag cagaagtaat tccagcagag acagggcaag aaacagcata cttcctctta	5100
aaattagcag gaagatggcc agtaaaaaca gtacatacag acaatggcag caatttcacc	5160
agtactacag ttaaggccgc ctgttggtgg gcggggatca agcaggaatt tggcattccc	5220
tacaatcccc aaagtcaagg agtaatagaa tctatgaata aagaattaaa gaaaattata	5280
ggacaggtaa gagatcaggc tgaacatctt aagacagcag taaaaatggc agtattcatc	5340
cacaatttta aaagaaaagg ggggattggg gggtagagtg caggggaaag aatagtagac	5400
ataatagcaa cagacataca aactaaagaa ttacaaaaac aaattacaaa aattcaaat	5460
tttcgggttt attacaggga cagcagagat ccagtttgga aaggaccagc aaagctcctc	5520
tggaaagggtg aaggggcagt agtaatacaa gataatagtg acataaaagt agtgccaaga	5580
agaaaagcaa agatcatcag ggattatgga aaacagatgg caggtgatga ttgtgtggca	5640
agtagacagg atgaggatta acacatggaa ttccggagcg gccgcaggag ctttgttcct	5700
tgggttcttg ggagcagcag gaagcactat gggcgacgcg tcaatgacgc tgacggtaca	5760
ggccagacaa ttattgtctg gtatagtgca gcagcagaac aatttgctga gggctattga	5820
ggcgcaacag catctgttgc aactcacagt ctggggcatc aagcagctcc aggcaagaat	5880
cctggctgtg gaaagatacc taaaggatca acagctcctg gggatttggg gttgctctgg	5940
aaaactcatt tgcaccactg ctgtgccttg gaatgctagt tggagtaata aatctctgga	6000
acagatttgg aatcacacga cctggatgga gtgggacaga gaaattaaca attacacaag	6060
cttcgcgga attcacccca ccagtgcagg ctgcctatca gaaagtgggtg gctgggtgtg	6120
ctaatagcct ggcccacaag tatcactaag ctgcgtttct tgctgtccaa tttctattaa	6180



aggttccttt gttccctaag tccaactact aaactggggg atattatgaa gggccttgag	6240
catctggatt ctgcctaata aaaaacattt attttcattg caatgatgta tttaaattat	6300
ttctgaatat ttactaaaa agggaatgtg ggaggtcagt gcatttaaaa cataaagaaa	6360
tgaagagcta gttcaaacct tgggaaaata cactatatct taaactccat gaaagaaggt	6420
gaggctgcaa acagctaattg cacattggca acagcccctg atgcctatgc cttattcatc	6480
cctcagaaaa ggattcaagt agaggcttga tttggagggt aaagttttgc tatgctgtat	6540
tttacattac ttattgtttt agctgtcctc atgaatgtct tttcactacc catttgctta	6600
tcctgcatct ctcagccttg actccactca gttctcttgc ttagagatac cacctttccc	6660
ctgaagtgtt ccttccatgt tttacggcga gatggtttct cctcgcttgg ccactcagcc	6720
ttagttgtct ctgttgtctt atagaggctt acttgaagaa ggaaaaacag ggggcatggt	6780
ttgactgtcc tgtgagccct tcttccctgc ctccccact cacagtgacc cggaatccct	6840
cgacatggca gtctagcact agtgcgccg cagatctgct tcctcgctca ctgactcgct	6900
gcgctcggtc gttcggctgc ggcgagcgg atcagctcac tcaaaggcgg taatacgggt	6960
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc	7020
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga	7080
gcacacaaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata	7140
ccaggcgttt cccctggaa gctccctcgt gcgctctcct gttccgacct tgccgcttac	7200
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg	7260
taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc	7320
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggttaag	7380
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatatg	7440
aggcgggtgct acagagttct tgaagtgggt gcctaactac ggctacacta gaagaacagt	7500
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg	7560
atccggcaaa caaaccaccg ctggtagcgg tgggtttttt gtttgcaagc agcagattac	7620
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacgggggt ctgacgctca	7680
gtggaacgaa aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac	7740
ctagatcctt tttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac	7800
ttgggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt	7860
tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggtt	7920
accatctggc ccagtgctg caatgatacc gcgagacca cgctcaccgg ctccagattt	7980
atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggctcctg caactttatc	8040

cgctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa 8100  
 tagtttgccg aacgttggtg ccattgctac aggcacgtg gtgtcacgct cgtcgtttgg 8160  
 tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat ccccatgtt 8220  
 gtgcaaaaaa gcggttagct ctttcggtcc tccgacgtt gtcagaagta agttggccgc 8280  
 agtgttatca ctcatgggta tggcagcact gcataattct cttactgtca tgccatccgt 8340  
 aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgatatgcg 8400  
 gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 8460  
 tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 8520  
 gctgttgaga tccagttcga tgtaaccac tcgtgcaccc aactgatctt cagcatcttt 8580  
 tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaatgccg caaaaaaggg 8640  
 aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag 8700  
 catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 8760  
 acaaataggg gttccgcgca catttccccg aaaagtgcc cctgacggga tcccctgagg 8820  
 gggcccccat gggctagagg atccggcctc ggcctctgca taaataaaaa aaattagtca 8880  
 gccatgagc 8889

<210> 7

<211> 4180

<212> DNA

<213> Artificial

<220>

<223> pLP2

<400> 7

aatgtagtct tatgcaatac tcttgtagtc ttgcaacatg gtaacgatga gttagcaaca 60  
 tgccttaciaa ggagagaaaa agcaccgtgc atgccgattg gtggaagtaa ggtggtacga 120  
 tcgtgcctta ttaggaaggc aacagacggg tctgacatgg attggacgaa ccactgaatt 180  
 ccgcattgca gagatattgt atttaagtgc ctagctcgat acaataaacg ccatttgacc 240  
 attcaccaca ttgggtgtgca cctccaagct cgagctcggt tagtgaaccg tcagatcgcc 300  
 tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 360  
 ccctcgaagc tagtcgatta ggcattctct atggcaggaa gaagcggaga cagcgacgaa 420

gacctcctca aggcagtcag actcatcaag tttctctatc aaagcaaccc acctcccaat	480
cccgagggga cccgacaggc ccgaaggaat agaagaagaa ggtggagaga gagacagaga	540
cagatccatt cgattagtga acggatcctt agcacttatc tgggacgatc tgccggagcct	600
gtgcctcttc agctaccacc gcttgagaga cttactcttg attgtaacga ggattgtgga	660
acttctggga cgcaggggggt ggggaagccct caaatattgg tggaatctcc tacaatattg	720
gagtcaggag ctaaagaata gtgctgttag cttgctcaat gccacagcta tagcagtagc	780
tgaggggaca gatagggtta tagaagtagt acaagaagct tggcactggc cgtcgtttta	840
caacgtcgtg atctgagcct gggagatctc tggctaacta gggaaaccac tgcttaagcc	900
tcaataaagc ttgccttgag tgcttcaagt agtgtgtgcc cgtctgttgt gtgactctgg	960
taactagaga tcaggaaaac cctggcgta cccaacttaa tcgccttgca gcacatcccc	1020
ctttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc	1080
gcagcctgaa tggcgaatgg cgcctgatgc ggtatcttct ccttacgcat ctgtgcggta	1140
tttcacaccg catacgtcaa agcaaccata gtacgcgccc tgtagcggcg cattaagcgc	1200
ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgcgc tagcgcgcgc	1260
tcctttcgct ttcttccctt cctttctcgc cacgttcgcc ggctttcccc gtcaagctct	1320
aaatcggggg ctcccttttag ggttcgatt tagtgcttta cggcacctcg accccaaaaa	1380
acttgatttg ggtgatggtt cacgtagtgg gccatcgccc tgatagacgg tttttcgccc	1440
tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact	1500
caaccctatc tcgggctatt cttttgattt ataagggatt ttgccgattt cggcctattg	1560
gttaaaaaat gagctgattt acaaaaaatt taacgcgaat ttaacaaaa tattaacgtt	1620
tacaatttta tgggtgactc tcagtacaat ctgctctgat gccgcatagt taagccagcc	1680
ccgacaccgg ccaacaccgg ctgacgcgcc ctgacgggct tgtctgctcc cggcatccgc	1740
ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt cagaggtttt caccgtcatc	1800
accgaaacgc gcgagacgaa agggcctcgt gatacgcta tttttatagg ttaatgtcat	1860
gataataatg gtttcttaga cgtcaggtgg cacttttcgg ggaaatgtgc gcggaacccc	1920
tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac aataaccctg	1980
ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc	2040
ccttattccc ttttttgagg cattttgcct tcctgttttt gctcaccag aaacgctgg	2100
gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg ggttacatcg aactggatct	2160
caacagcggg aagatccttg agagttttcg ccccgaagaa cgttttccaa tgatgagcac	2220
ttttaaggtt ctgctatgtg gcgcgggtatt atcccgtatt gacgccgggc aagagcaact	2280

cggtcgccgc atacactatt ctcagaatga cttgggttgag tactcaccag tcacagaaaa	2340
gcatcttacg gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga	2400
taacactgcg gccaaacttac ttctgacaac gatcggagga ccgaaggagc taaccgcttt	2460
tttgcacaa atgggggatc atgtaactcg ccttgatcgt tgggaaccgg agctgaatga	2520
agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa caacgttgcg	2580
caaactatta actggcgaac tacttactct agcttcccgg caacaattaa tagactggat	2640
ggaggcggat aaagttgcag gaccacttct gcgctcggcc cttccggctg gctggtttat	2700
tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt atcattgcag cactggggcc	2760
agatggtaag ccctcccgtc tcgtagtatt ctacacgacg gggagtcagg caactatgga	2820
tgaacgaaat agacagatcg ctgagatagg tgcctcactg attaagcatt ggtaactgtc	2880
agaccaagtt tactcatata tacttttagat tgatttaaaa cttcattttt aatttaaaag	2940
gatctagggtg aagatccttt ttgataatct catgacccaa atcccttaac gtgagttttc	3000
gttcactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt	3060
tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt	3120
gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat	3180
accaaatact gttcttctag tgtagccgta gttaggccac cacttcaaga actctgtagc	3240
accgcctaca tacctcgctc tgctaatacct gttaccagtg gctgctgcca gtggcgataa	3300
gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg	3360
ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag	3420
atacctacag cgtgagctat gagaaagcgc cacgcttccc gaaggagaaa aggcggacag	3480
gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa	3540
cgcctgggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt	3600
gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg	3660
gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc	3720
tgtggataac cgtattaccg cctttgagtg agctgatacc gctcgcgcga gccgaacgac	3780
cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct	3840
ccccgcgcgt tggccgattc attaatgcag ctggcacgac aggtttcccg actggaaagc	3900
gggcagtgag cgcaacgcaa ttaatgtgag ttagctcact cattaggcac ccaggcttt	3960
acactttatg cttccggctc gtatgttggtg tggaattgtg agcggataac aatttcacac	4020
aggaaacagc tatgacatga ttacgaattc gatgtacggg ccagatatat gcgtatctga	4080

ggggactagg gtgtgttttag gcgaaaagcg gggcttcggt tgtacgcggt taggagtccc 4140  
 ctcaggatat agtagtttcg cttttgcata gggaggggga 4180

<210> 8

<211> 5821

<212> DNA

<213> Artificial

<220>

<223> pLP/VSVG

<400> 8

ttggcccatt gcatacgttg tatccatata ataatatgta catttatatt ggctcatgtc 60  
 caacattacc gccatgttga cattgattat tgactagtta ttaatagtaa tcaattacgg 120  
 ggtcattagt tcatagccca tatatggagt tccgcgttac ataacttacg gtaaattggcc 180  
 cgcttggtg accgcccac gacccccgc cattgacgtc aataatgacg tatgttccca 240  
 tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg 300  
 cccacttggc agtacatcaa gtgtatcata tgccaagtac gcccctatt gacgtcaatg 360  
 acggtaaatg gcccgcttg cattatgcc agtacatgac cttatgggac tttcctactt 420  
 ggcagtacat ctacgtatta gtcacgcta ttaccatggt gatgcgggtt tggcagtaca 480  
 tcaatgggag tggatagcgg tttgactcac ggggatttcc aagtctccac ccattgacg 540  
 tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt cgtaacaact 600  
 ccgccccatt gacgcaaatg ggcggtaggc gtgtacggtg ggaggtctat ataagcagag 660  
 ctcgttttagt gaaccgtcag atcgcttga gacgccatcc acgctgtttt gacctccata 720  
 gaagacaccg ggaccgatcc agcctcccct cgaagcttac atgtggtacc gagctcggat 780  
 cctgagaact tcagggtgag tctatgggac ccttgatgtt ttctttcccc ttcttttcta 840  
 tgggttaagtt catgtcatag gaaggggaga agtaacaggg tacacatatt gaccaaata 900  
 gggtaatttt gcatttgtaa ttttaaaaaa tgctttcttc ttttaataata cttttttgtt 960  
 tatcttattt ctaatacttt ccctaattct tttctttcag ggcaataatg atacaatgta 1020  
 tcatgcctct ttgcaccatt ctaaagaata acagtataa tttctgggtt aaggcaatag 1080  
 caatatttct gcatataaat atttctgcat ataaattgta actgatgtaa gaggtttcat 1140  
 attgctaata gcagctacaa tccagctacc attctgcttt tattttatgg ttgggataag 1200  
 gctggattat tctgagtcca agctaggccc ttttgctaata catgttcata cctcttatct 1260

tcctcccaca gctcctgggc aacgtgctgg tctgtgtgct ggcccatcac tttggcaaag	1320
cacgtgagat ctgaattctg acactatgaa gtgccttttg tacttagcct ttttattcat	1380
tgggggtgaat tgcaagttca ccatagtttt tccacacaac caaaaaggaa actggaaaaa	1440
tgttccttct aattaccatt attgcccgtc aagctcagat tttaaattggc ataatgactt	1500
aataggcaca gccttacaag tcaaaatgcc caagagtcac aaggctattc aagcagacgg	1560
ttggatgtgt catgcttcca aatgggtcac tacttgtgat ttccgctggg atggaccgaa	1620
gtatataaca cattccatcc gatccttcac tccatctgta gaacaatgca aggaaagcat	1680
tgaacaaacg aaacaaggaa cttggctgaa tccaggcttc cctcctcaa gttgtggata	1740
tgcaactgtg acggatgccg aagcagtgat tgtccagggt actcctcacc atgtgctggg	1800
tgatgaatac acaggagaat gggttgattc acagttcatc aacggaaaat gcagcaatta	1860
catatgcccc actgtccata actctacaac ctggcattct gactataagg tcaaagggct	1920
atgtgattct aacctcattt ccatggacat caccttcttc tcagaggacg gagagctatc	1980
atccctggga aaggagggca cagggttcag aagtaactac tttgcttatg aaactggagg	2040
caaggcctgc aaaatgcaat actgcaagca ttggggagtc agactcccat cagggtgtctg	2100
gttcgagatg gctgataagg atctcttttg tgcagccaga ttccctgaat gcccagaagg	2160
gtcaagtatc tctgctccat ctcagacctc agtggatgta agtctaattc aggacgttga	2220
gaggatcttg gattattccc tctgccaaga aacctggagc aaaatcagag cgggtcttcc	2280
aatctctcca gtggatctca gctatcttgc tcctaaaaac ccaggaaccg gtcctgcttt	2340
caccataatc aatggtaccc taaaatactt tgagaccaga tacatcagag tcgatattgc	2400
tgctccaatc ctctcaagaa tggtcggaat gatcagtgga actaccacag aaagggaaact	2460
gtgggatgac tgggcaccat atgaagacgt ggaaattgga cccaatggag ttctgaggac	2520
cagttcagga tataagtttc ctttatacat gattggacat ggtatgttgg actccgatct	2580
tcatcttagc tcaaaggctc aggtgttcga acatcctcac attcaagacg ctgcttcgca	2640
acttcctgat gatgagagtt tatttttttg tgatactggg ctatccaaaa atccaatcga	2700
gcttgtagaa ggttggttca gtagttggaa aagctctatt gcctcttttt tctttatcat	2760
agggttaatc attggactat tcttggttct ccgagttggg atccatcttt gcattaaatt	2820
aaagcacacc aagaaaagac agatttatac agacatagag atgaaccgac ttggaaagta	2880
actcaaatec tgcacaacag attcttcatg tttggaccaa atcaacttgt gataccatgc	2940
tcaaagaggc ctcaattata tttgagtttt taatttttat gaaaaaaaaa aaaaaaacg	3000
gaattcacc caccagtgc ggctgcctat cagaaagtgg tggctgggtg ggctaagcc	3060

ctggcccaca agtatcacta agctcgcttt cttgctgtcc aatttctatt aaagggttcct	3120
ttgttcccta agtccaacta ctaaactggg ggatattatg aagggccttg agcatctgga	3180
ttctgcctaa taaaaaacat ttattttcat tgcaatgatg tatttaaatt atttctgaat	3240
attttactaa aaagggaatg tgggaggtca gtgcatttaa aacataaaga aatgaagagc	3300
tagttcaaac cttgggaaaa tacactatat cttaaactcc atgaaagaag gtgaggctgc	3360
aaacagctaa tgcacattgg caacagcccc tgatgcctat gccttattca tccctcagaa	3420
aaggattcaa gtagaggctt gatttggagg ttaaagtttt gctatgctgt attttacatt	3480
acttattgtt ttagctgtcc tcatgaatgt cttttcacta cccatttgct tatcctgcat	3540
ctctcagcct tgactccact cagttctctt gcttagagat accaccttcc cctgaagtg	3600
ttccttccat gttttacggc gagatggttt ctctcgcct ggccactcag ccttagttgt	3660
ctctgttgtc ttatagaggt ctacttgaag aaggaaaaac agggggcatg gtttgactgt	3720
cctgtgagcc cttcttccct gcctcccca ctacagtgga cccggaatcc ctcgacatgg	3780
cagtctagca ctagtgcggc cgcagatctg cttctcgtc cactgactcg ctgcgctcgg	3840
tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag	3900
aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc	3960
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca	4020
aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt	4080
ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc	4140
tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc	4200
tcagttcggg gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc	4260
ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caaccggta agacacgact	4320
tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg	4380
ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggtg	4440
tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca	4500
aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa	4560
aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg	4620
aaaactcacg ttaagggatt ttgggtcatga gattatcaaa aaggatcttc acctagatcc	4680
ttttaaatta aaaatgaagt tttaaataca tctaaagtat atatgagtaa acttggtctg	4740
acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat	4800
ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg	4860
gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa	4920

```

taaaccagcc agccggaagg gccgagcgca gaagtgggtcc tgcaacttta tccgcctcca 4980
tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc 5040
gcaacgttgt tgccattgct acaggcatcg tgggtgtcacg ctcgtegttt ggtatggctt 5100
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 5160
aagcgggttag ctcttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 5220
cactcatggc tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 5280
tttctgtgac tgggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 5340
gttgctcttg cccggcgtca atacgggata ataccgcgcc acatagcaga actttaaaag 5400
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 5460
gatccagttc gatgtaacct actcgtgcac ccaactgatc ttcagcatct tttactttca 5520
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 5580
cgacacggaa atgttgaata ctcatactct tcctttttca atattattga agcatttatc 5640
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 5700
gggttccgag cacatttccc cgaaaagtgc cacctgacgg gatccctga gggggccccc 5760
atgggctaga ggatccggcc tcggcctctg cataaataaa aaaaattagt cagccatgag 5820
c 5821

```

<210> 9

<211> 47

<212> DNA

<213> Artificial

<220>

<223> lamin A/C control oligo

<400> 9

caccgtgttc ttctggaagt ccagcgaact ggacttccag aagaaca 47

<210> 10

<211> 47

<212> DNA

<213> Artificial



<220>

<223> lamin A/C control oligo

<400> 10

aaaatgttct tctggaagtc cagttcgctg gacttccaga agaacac

47

<210> 11

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Sequencing primer

<400> 11

ggactatcat atgcttaccg

20

<210> 12

<211> 17

<212> DNA

<213> Artificial

<220>

<223> Sequencing primer

<400> 12

caggaaacag ctatgac

17

<210> 13

<211> 20

<212> DNA

<213> Artificial

<220>

<223> U6 promoter sequence

<400> 13

aaggtcgggc aggaagaggg

20

<210> 14

<211> 20

<212> DNA

<213> Artificial

<220>

<223> U6 promoter sequence

<400> 14

agcgagcacg gtgtttcgtc

20

<210> 15

<211> 29

<212> DNA

<213> Artificial

<220>

<223> U6 promoter sequence with Asp718 and Not I at 5' end

<400> 15

gtgggtacca aggtcgggca ggaagaggg

29

<210> 16

<211> 33

<212> DNA

<213> Artificial

<220>

<223> U6 promoter sequence with Asp718 and Not I at 5' end

<400> 16

gtggcgccg cggtgtttcg tcctttccac aag

33

<210> 17

<211> 44

<212> DNA

<213> Artificial

<220>

<223> primer for ccdB gene

<400> 17

gtggcgggccg caaagatcct ccagtggatc cggcttacta aaag

44

<210> 18

<211> 57

<212> DNA

<213> Artificial

<220>

<223> primer for ccdB gene

<400> 18

gtgctcgaga aaaaagtcga cacggagccc tccagttata ttccccagaa catcagg

57

<210> 19

<211> 30

<212> DNA

<213> Artificial

<220>

<223> part of double stranded oligo containing BsaI and NotI site to engineer BsaI vector

<400> 19

gagaccgcgg ccgcttctcg aggtctcatt

30

<210> 20

<211> 30

<212> DNA

<213> Artificial

<220>

<223> part of double stranded oligo containing BsaI and NotI site to engineer BsaI vector

<400> 20

tgagacctcg agaagcggcc gcggtctccg

30

<210> 21

<211> 31

<212> DNA

<213> Artificial

<220>

<223> Primer for new ccdB region with NotI site, BsaI site, XbaI site

<400> 21

cacgcggccg ctggatccgg cttactaaaa g

31

<210> 22

<211> 43

<212> DNA

<213> Artificial

<220>

<223> Primer for new ccdB region

<400> 22

cactctagaa aaaatgagac cttatattcc ccagaacatc agg

43

<210> 23

<211> 7

<212> PRT

<213> Artificial

<220>

<223> Consensus cleavage site

<220>

<221> UNSURE

<222> (2) .. (3)

<223> Xaa can be any amino acid

<220>

<221> UNSURE

<222> (5) .. (5)

<223> Xaa can be any amino acid.

<220>

<221> UNSURE

<222> (7) .. (7)

<223> Xaa can be any amino acid except Proline.

<400> 23

Glu Xaa Xaa Tyr Xaa Gln Xaa  
1 5

<210> 24

<211> 7

<212> PRT

<213> Artificial

<220>

<223> TEV1 cleavage site

<220>

<221> UNSURE

<222> (7) .. (7)

<223> Xaa can be any amino acid except Proline.

<400> 24

Glu Asn Leu Tyr Phe Gln Xaa  
1 5

<210> 25

<211> 7

<212> PRT

<213> Artificial

<220>

<223> TEV 2 cleavage site

<220>

<221> UNSURE

<222> (7) .. (7)

<223> Xaa can be any amino acid except Proline.

<400> 25

Glu Thr Leu Tyr Ile Gln Xaa  
1 5

<210> 26

<211> 5

<212> PRT

<213> Artificial

<220>

<223> EK Cleavage site

<400> 26

Asp Asp Asp Asp Lys  
1 5

<210> 27

<211> 7

<212> PRT

<213> Artificial

<220>

<223> TEV1 cleavage site (cleaves between Gln and Gly)

<400> 27

Glu Asn Leu Tyr Phe Gln Gly  
1 5

<210> 28

<211> 3

<212> PRT

<213> Artificial

<220>

<223> ulp1 protease recognition site

<400> 28

Gly Gly Ser  
1

<210> 29

<211> 5

<212> PRT

<213> Artificial

<220>

<223> Peptide tag

<400> 29

Phe His His Thr Thr  
1 5

<210> 30

<211> 6

<212> PRT

<213> Artificial

<220>

<223> FlAsH tags

<220>

<221> UNSURE

<222> (3)..(3)

<223> Xaa can be any amino acid. In many instances, Xaa is an amino acid with high a-helical propensity

<220>

<221> UNSURE

<222> (4)..(4)

<223> Xaa can be any amino acid. In many instances, Xaa is an amino acid with high a-helical propensity

<400> 30

Cys Cys Xaa Xaa Cys Cys  
1 5

<210> 31

<211> 32

<212> DNA

<213> Artificial

<220>

<223> BsaI digestion site

<400> 31  
acaccggaga ccggtctcat tttttttcta ga

32

<210> 32

<211> 32

<212> DNA

<213> Artificial

<220>

<223> BsaI digestion site complementary sequence

<400> 32  
gctagaaaaa aaatgagacc ggtctccggt gt

32



<210> 33

<211> 13

<212> DNA

<213> Artificial

<220>

<223> BsaI digestion fragment

<400> 33  
 ttttttttct aga

13

<210> 34

<211> 24

<212> DNA

<213> Artificial

<220>

<223> possible insert with overhang into pENTR/U6-BsaI-ccdB

<220>

<221> Unsure

<222> (6) .. (24)

<223> N can be any nucleotide.

<400> 34  
 caccgnnnnn nnnnnnnnnn nnnn

24

<210> 35

<211> 24

<212> DNA

<213> Artificial

<220>

<223> possible insert (complementary seq. with overhang) into pENTR/U6-BsaI-ccd

<220>

<221> Unsure

<222> (5)..(23)

<223> N can be any nucleotide.

<400> 35

aaaaaaaaaaaaa nnnnnnnnnnn nnnnc

24

<210> 36

<211> 21

<212> DNA

<213> Artificial

<220>

<223> siRNA core molecule with overhang

<400> 36

uucagugagu agagucuat t

21

<210> 37

<211> 21

<212> DNA

<213> Artificial

<220>

<223> siRNA core molecule complementary seq. with overhang

<400> 37

uauagactcta ctcacugaat t

21

<210> 38

<211> 34

<212> RNA

<213> Artificial

<220>

<223> miRNA sequence with mismatches

<400> 38

gcgacuguaa acauccucga cuggaagcug ugaa

34

<210> 39

<211> 37

<212> RNA

<213> Artificial

<220>

<223> miRNA complementary sequence with mismatches

<400> 39

gccacagaug ggcuuucagu cgguaguug cagcugc

37

<210> 40

<211> 24

<212> RNA

<213> Artificial

<220>

<223> diced miRNA with mismatches

<400> 40

uguaaacauc cucgacugga agcu

24

<210> 41

<211> 22

<212> RNA

<213> Artificial

<220>

<223> diced miRNA complementary sequence with mismatches

<400> 41

cuuucagucg gauguuugca gc

22

<210> 42

<211> 23

<212> RNA

<213> Artificial

<220>

<223> siRNA

<400> 42

auuucgaagu auuccgcgua cgu

23

<210> 43

<211> 23

<212> RNA

<213> Artificial

<220>

<223> siRNA complementary sequence

<400> 43

acguacgcgg aaacuucga aau

23

<210> 44

<211> 25

<212> RNA

<213> Artificial

<220>

<223> shRNA

<400> 44

auuucgaagu auuccgcgua cguuu

25

<210> 45

<211> 28

<212> RNA

<213> Artificial

<220>

<223> shRNA complementary sequence

<400> 45

cgacguacgc ggaauacuuc gaaauuuu

28

<210> 46

<211> 41

<212> RNA

<213> Artificial

<220>

<223> miRNA with mismatches

<400> 46

cucgagaucu gcgccguacg cggaauacuu cgaaauguga a

41

<210> 47

<211> 47

<212> RNA

<213> Artificial

<220>

<223> miRNA complementary sequence with mismatches

<400> 47

gccacagaug auuucgaagu auuccgcgua cguugcggau ccucgag

47

<210> 48

<211> 23

<212> DNA

<213> Artificial

<220>

<223> Directional TOPO cloning site with gene of interest

<220>

<221> Unsure

<222> (13)..(18)

<223> N can be any nucleotide.

<400> 48

cccttcacca tgnnnnnnaa ggg

23

<210> 49

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Complementary sequence of Directional TOPO cloning site with gene  
of interes

<220>

<221> Unsure

<222> (6)..(11)

<223> N can be any nucleotide.

<400> 49

cccttnnnnn ncatggtggg tgaaggg

27

<210> 50

<211> 11

<212> DNA

<213> Artificial

<220>

<223> TOPO cloning site

<400> 50

cccttaaggg c

11

<210> 51

<211> 15

<212> DNA

<213> Artificial

<220>

<223> TOPO cloning site complementary sequence with overhang

<400> 51

gcccttggtg aaggg

15